PCT
WORLD INTELLECTUAL
Internati

WO 9602563A1

INTERNATIONAL APPLICATION PUBLISHED U

(51) International Patent Classification 6 :
C07K 1/00, C12Q 1/70, C12P 21/06, C12N 5/00, C07H 21/04

A1

(43) International Publication Date: 1 February 1996 (01.02.96)

(21) International Application Number: PCT/US95/08700

(22) International Filing Date: 13 July 1995 (13.07.95)

(30) Priority Data:
08/275,614 13 July 1994 (13.07.94) US

(71) Applicant: CORNELL RESEARCH FOUNDATION, INC. [US/US]; Suite 105, 20 Thornwood Drive, Ithaca, NY 14850 (US).

(72) Inventor: O'DONNELL, Michael, E.; 16 Maple Lane, Hastings on Hudson, NY 10706 (US).

(74) Agents: GOLDMAN, Michael, L. et al.; Nixon, Hargrave, Devans & Doyle, Clinton Square, P.O. Box 1051, Rochester, NY 14603 (US).

(81) Designated States: JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

**Published**
*With international search report.*

(54) Title: EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN 1 PROTEIN AND ITS EXPRESSION AND RECOVERY

(57) Abstract

A process for expressing and recovering Epstein-Barr nuclear antigen 1 (EBNA1) protein or polypeptide treats cells having a nucleus containing expressed EBNA1 protein or polypeptide to recover the nucleus containing the expressed EBNA1 protein or polypeptide. The nucleus containing the expressed EBNA1 protein or polypeptide is then separated into a liquid fraction containing the expressed EBNA1 protein or polypeptide and a solid fraction containing substantially all DNA from the nucleus. The liquid fraction is separated from the solid fraction, and EBNA1 protein or polypeptide is recovered from the liquid fraction. Also encompassed by the present invention is an EBNA1 protein or polypeptide having substantially no components which generate false positive readings when used to detect Epstein-Barr virus in human serum, the DNA molecule encoding it, and recombinant expression of the protein. The protein is useful in a method for detection of Epstein-Barr virus.

## *FOR THE PURPOSES OF INFORMATION ONLY*

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

| | |
|---|---|
| AT | Austria |
| AU | Australia |
| BB | Barbados |
| BE | Belgium |
| BF | Burkina Faso |
| BG | Bulgaria |
| BJ | Benin |
| BR | Brazil |
| BY | Belarus |
| CA | Canada |
| CF | Central African Republic |
| CG | Congo |
| CH | Switzerland |
| CI | Côte d'Ivoire |
| CM | Cameroon |
| CN | China |
| CS | Czechoslovakia |
| CZ | Czech Republic |
| DE | Germany |
| DK | Denmark |
| ES | Spain |
| FI | Finland |
| FR | France |
| GA | Gabon |
| GB | United Kingdom |
| GE | Georgia |
| GN | Guinea |
| GR | Greece |
| HU | Hungary |
| IE | Ireland |
| IT | Italy |
| JP | Japan |
| KE | Kenya |
| KG | Kyrgystan |
| KP | Democratic People's Republic of Korea |
| KR | Republic of Korea |
| KZ | Kazakhstan |
| LI | Liechtenstein |
| LK | Sri Lanka |
| LU | Luxembourg |
| LV | Latvia |
| MC | Monaco |
| MD | Republic of Moldova |
| MG | Madagascar |
| ML | Mali |
| MN | Mongolia |
| MR | Mauritania |
| MW | Malawi |
| NE | Niger |
| NL | Netherlands |
| NO | Norway |
| NZ | New Zealand |
| PL | Poland |
| PT | Portugal |
| RO | Romania |
| RU | Russian Federation |
| SD | Sudan |
| SE | Sweden |
| SI | Slovenia |
| SK | Slovakia |
| SN | Senegal |
| TD | Chad |
| TG | Togo |
| TJ | Tajikistan |
| TT | Trinidad and Tobago |
| UA | Ukraine |
| US | United States of America |
| UZ | Uzbekistan |
| VN | Viet Nam |

# EPSTEIN-BARR VIRUS NUCLEAR ANTIGEN 1 PROTEIN AND ITS EXPRESSION AND RECOVERY

The subject matter of this invention was developed with the support of the United States Government (NIH Grant Nos. ROI-GM38839 and ROI-CA531-01).

## FIELD OF THE INVENTION

The present invention relates to Epstein-Barr virus nuclear antigen 1 (EBNA1) protein and its expression and recovery. More particularly, the present invention relates to a process for recovering EBNA1 protein or polypeptide from cells having a nucleus containing expressed EBNA1 protein or polypeptide.

## BACKGROUND

Epstein-Barr virus ("EBV"), a human herpesvirus, is one of the most common viruses infecting man, and antibodies to EBV proteins are present in greater than 80% of human serum samples. Milman et al., "Carboxyl-terminal domain of the Epstein-Barr virus nuclear antigen is highly immunogenic in man," Proc. Natl. Acad. Sci. USA, 82:6300-04 (1985), which is hereby incorporated by reference.

EBV was discovered during the course of attempts to learn the cause of lymphoma that was the most common tumor affecting children in certain parts of East Africa. The clinical syndrome, which was described in detail by Dennis Burkitt in 1958, had, in retrospect, been known to clinicians and pathologists since the beginning of the 20th century. However, through Burkitt's efforts, the disease was unified into a clearly delineated entity with characteristic clinical, pathological, and epidemiological features. See Burkitt D., "A sarcoma involving the jaws in

African children," Br. J. Surg., 46:218-223 (1958), which is hereby incorporated by reference. In 1964, Epstein and Barr reported the first successful attempt to establish continuous lymphoblastoid cell lines from explants of Burkitt's lymphoma ("BL"), which were eventually found to be infected with EBV by W. and G. Henle in 1966. Epstein et al., "Cultivation *in vitro* of human lymphoblasts from Burkitt's malignant lymphoma," Lancet, 1:252-53 (1964) and Henle et al., "Immunofluorescence in cells derived from Burkitt's lymphoma," J. Bacteriol., 91:1248-1256 (1966), which are hereby incorporated by reference.

In addition to its involvement in BL, EBV is the etiological agent of infectious mononucleosis and has been implicated in the pathogenesis of nasopharyngeal carcinoma. EBV can also induce fatal lymphoproliferative disease, sometimes with the features of frank lymphoma, in certain patients with global immunodeficiency that is either congenital (such as severe combined immunodeficiency or ataxia telangiectasia) or acquired as the result of immunosuppression for organ or tissue transplantation or due to AIDS. Kieff et al., "Epstein-Barr Virus and Its Replication," Chapter 67, pp. 1889-1920 and Miller, "Epstein-Barr Virus: Biology, Pathogenesis, and Medical Aspects," Chapter 68, pp. 1921-1958, in Virology, Second Edition, edited by B. N. Fields, D. M. Knipe et al., Raven Press, Ltd., New York, 1990, which are hereby incorporated by reference.

It has been determined that the principal biological activity of EBV that underlies its role in the pathogenesis of lymphoproliferative diseases is the ability of the virus to cause indefinite *in vitro* proliferation of lymphocytes, a process termed "immortalization." The sequence of several events in the process of immortalization has been defined. The process is thought to consist of two phases: (i) an initial phase of B-cell activation, triggered

by virus binding to the cell surface, and (ii) a subsequent phase of permanent blastogenesis which requires the expression of 10 EBV gene-encoded products - the primary of which is EBV nuclear antigen 1 ("EBNA1").

The 172,000-base-pair ("bp") DNA genome of EBV is found in all "immortalized" permanent B-cell lymphoblast lines as multicopy latent extrachromosomal circular DNA plasmids or episomes. Only EBNA1 is essential for the replication of these EBV plasmids. The EBNA1 protein comprises 641 amino acids ("aa"). One-third of EBNA1 (aa 90 to 325) consists of a repetitive array of glycine ("Gly") and alanine ("Ala") amino acid residues. Shah et al., "Binding of EBNA-1 to DNA Creates Protease-Resistant Domain That Encompasses the DNA Recognition and Dimerization Functions," Journal of Virology, 66:6:3355-62 (1992) and Yates et al., "Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells," Nature, 318:812-15 (1985), which are hereby incorporated by reference. The size of the repeat array varies among different EBV isolates and the EBNA1 polypeptide shows corresponding size variations ranging from 68 kDa to 84 kDa. Milman et al., "Carboxyl-terminal domain of the Epstein-Barr virus nuclear antigen is highly immunogenic in man," Proc. Natl. Acad. Sci. USA, 82:6300-04 (1985), which is hereby incorporated by reference.

Additionally, it has been reported that the Gly-Ala repeat sequence has homology to cellular DNA, and antisera to Gly-Ala repeat-containing peptides also react with cellular proteins, e.g., *E. coli*, mammalian or baculovirus cellular proteins containing glycine plus alanine-rich regions. Id.

EBNA1 protein binds *in trans* to the latent origin of replication, *oriP*, at multiple sites present in the two regions of *oriP* which were found to be necessary and sufficient for origin function. One of these regions is

composed of 20 tandem copies of a 30-bp sequence (i.e., family of repeats), each of which contains an EBNA1 binding site. The other region includes four EBNA1 binding sites (dyad symmetry element), two of which are located within a 65-bp region of dyad symmetry. The interaction of EBNA1 with *oriP* occurs mainly through the carboxyl-terminal third of the protein.

It has been theorized that EBNA1 activates *oriP* to function not only as an origin of replication but also as a plasmid maintenance element and a transcriptional enhancer. Frappier et al., "Overproduction, Purification, and Characterization of EBNA1, the Origin Binding Protein of Epstein-Barr Virus," The Journal of Biological Chemistry, 266(12):7819-26, (1991), and Yates et al., "Dissection of DNA Replication and Enhancer Activation Functions of Epstein-Barr Virus Nuclear Antigen 1," Cancer Cells 6/Eukaryotic DNA Replication, pp. 197-205, Cold Spring Harbor Laboratory, 1988, which are hereby incorporated by reference.

Unfortunately, the production of biochemical assays to analyze the mechanism by which EBNA1 activates *oriP* to function as the origin of replication, a plasmid maintenance element, and a transcriptional enhancer has been difficult due to the lack of efficient systems for production of the virus and the very low amounts of gene products in transformed cells. Further, low protein expression has hindered the application of cell-derived EBNA1 protein as an antigen in a detection immunoassay for EBV.

Initially, researchers utilized *E. coli*-based expression systems in an attempt to produce the EBNA1 protein. For example, Orlowski et al., "Inhibition of Specific Binding of EBNA1 to DNA by Murine Monoclonal and Certain Human Polyclonal Antibodies," Virology, 176:638-42

(1992), expressed EBNA1 as a non-fusion protein in *E. coli* under control of the *lac* promoter.

Milman et al., "Carboxyl-terminal domain of the Epstein-Barr virus nuclear antigen is highly immunogenic in man," Proc. Natl. Acad. Sci. USA, 82:6300-04 (1985), synthesized the carboxyl-terminal one-third of EBNA1 encoded by the *Bam*HI restriction fragment K in *E. coli* by use of the expression plasmid pHE6. Expression of the EBNA1 fusion polypeptide was poor, i.e., only approximately 1.3 $\mu$g was recovered.

In Chen et al., "Separation of the Complex DNA Binding Domain of EBNA-1 into DNA Recognition and Dimerization Subdomains of Novel Structure," Journal of Virology, 67:8:4875-85 (1993), and Shah et al., "Binding of EBNA-1 to DNA Creates Protease-Resistant Domain That Encompasses the DNA Recognition and Dimerization Functions," Journal of Virology, 66:6:3355-62 (1992), the DNA binding and dimerization functions of EBNA1 were studied by creating a series of deletions and point mutations in the region of that protein spanning amino acids 408 to 641. Genes encoding for these modified forms of EBNA1 were cloned into plasmids and transformed into *E. coli*. Expression was poor in both studies.

The use of mammalian cell expression systems has also been described. For example, Yates et al., "Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells," Nature, 318:812-15 (1985), studied the functions of various segments of the gene encoding EBNA1 by deletion analysis. As shown in FIG. 5 of Yates et al., several of these deletions involve removal of the Gly-Ala repeat amino acid sequence. The genes encoding for these proteins were cloned into plasmids which were used to transfect human cells.

Middleton et al., "EBNA1 Can Link the Enhancer Element to the Initiator Element of the Epstein-Barr Virus

Plasmid Origin of DNA Replication," Journal of Virology, 66(1):489-95 (1992), which is hereby incorporated by reference, expressed EBNA1 in CV-1p cells by using an infectious simian virus (SV) 40 vector containing the EBNA1 gene. Expression was quite poor.

Hammarskjöld et al., "High-level expression of the Epstein-Barr virus EBNA1 protein in CV1 cells and human lymphoid cells using a SV40 late replacement vector," Gene, 43:41-50 (1986), which is hereby incorporated by reference, inserted the EBNA1 gene-containing EBV *Bam*HI-K fragment (B95-8 strain) into an expression vector composed of SV40 and pBR322 DNA. The vector was transfected into CV1 monkey cells and yielded EBNA1 protein (which included the entire Gly-Ala repeat unit) in 40-50% of the transfected cells.

Unfortunately, protein contaminants produced by current *E. coli* or mammalian cell expression systems can contribute to false positive readings when *E. coli* or mammalian cell-derived EBNA1 protein is used as an antigen in a detection immunoassay for EBV. This is due to crossreactivity of the detecting antibodies with *E. coli* or mammalian cellular contaminant proteins containing glycine plus alanine-rich regions, i.e., the antibodies could bind with the Gly-Ala repeat portion of the contaminant proteins thereby incorrectly indicating the presence of EBV. See Milman et al., "Carboxyl-terminal domain of the Epstein-Barr virus nuclear antigen is highly immunogenic in man," Proc. Natl. Acad. Sci. USA, 82:6300-04 (1985), which is hereby incorporated by reference.

Several researchers have attempted to express EBNA1 protein utilizing baculovirus expression systems, but with limited success with regard to quantity and purity of recovered protein. For example, Hearing et al., "Interaction of Epstein-Barr Virus Nuclear Antigen 1 with the Viral Latent Origin of Replication," Journal of Virology, 66(2):694-705 (1992), which is hereby incorporated

by reference, described a process for expression and purification of EBNA1 using a baculovirus expression system. The gene encoding the protein contained the entire EBNA1 open reading frame of the B95-8 virus isolate, including the entire Gly-Ala repeat amino acid sequence. As in the *E. coli* and mammalian cell expression systems discussed above, the presence of the Gly-Ala repeat in baculovirus cell-derived EBNA1 protein could contribute to false positive readings when such an EBNA1 protein is used as an antigen in a detection immunoassay for EBV. Hearing et al. also achieved poor expression levels.

Frappier et al., "Overproduction, Purification, and Characterization of EBNA1, the Origin Binding Protein of Epstein-Barr Virus," The Journal of Biological Chemistry, 266:12:7819-26 (1991), which is hereby incorporated by reference, expressed and purified the EBNA1 protein using a baculovirus expression system. In this process, a portion of the EBNA1 protein was expressed with an undefined deletion of amino acids in its Gly-Ala repeat amino acid sequence. As a result, due to the reference's failure to define the nature of the Gly-Ala repeat amino acid sequence deletion, its work cannot be repeated. Approximately 1.4 mg (representing a yield of 33%) of homogeneous 50-kDa baculovirus-derived EBNA1 protein was recovered. In addition, only a low yield of EBNA1 was recovered, because, after disruption of the nucleus, the EBNA1 protein and nuclear DNA were not adequately separated. Instead, there was a substantial gelatinous fraction containing both the protein and DNA. The EBNA1 protein in this fraction could not be recovered.

Since the above-described expression systems only teach how to produce small amounts of relatively impure EBNA1 protein, the use of the EBNA1 protein continues to be hampered. Further, the currently expressed EBNA1 proteins are susceptible to false positive readings when used as an

antigen in a detection immunoassay for EBV. There thus remains a need to achieve improved EBNA1 protein expression.

## SUMMARY OF THE INVENTION

One aspect of the present invention relates to a process for recovering EBNA1 protein or polypeptide. In this process, cells having a nucleus containing expressed EBNA1 protein or polypeptide are treated to recover the nucleus containing the expressed EBNA1 protein or polypeptide. The nucleus containing the expressed EBNA1 protein or polypeptide is then separated into a liquid fraction containing the expressed EBNA1 protein or polypeptide and a solid fraction containing substantially all DNA from the nucleus. The liquid fraction is separated from the solid fraction, and EBNA1 protein or polypeptide is recovered from the liquid fraction. This process produces abundant quantities of purified EBNA1 protein or polypeptide useful for diagnosis of EBV.

The present invention also relates to an isolated EBNA1 protein or polypeptide formulation having substantially no components which generate false positive readings when used to detect EBV in human serum. This isolated EBNA1 protein or polypeptide formulation can be utilized for detection of EBV in a sample of human tissue or body fluids. This detection process involves providing the isolated EBNA1 protein or polypeptide formulation as an antigen, contacting the sample with the antigen, and detecting any reaction which indicates that EBV is present in the sample using an assay system.

Additionally, the present invention provides an isolated DNA molecule encoding EBNA1 protein or polypeptide, a recombinant DNA expression system comprising an expression vector into which is inserted a heterologous DNA molecule encoding EBNA1 protein or polypeptide, and a host cell

incorporating a heterologous DNA molecule encoding EBNA1 protein or polypeptide, all of which have substantially no components which generate false positive readings when used to detect EBV in human serum.

The present invention also provides a process of expressing an EBNA1 protein coding sequence in a cell. In this process, an EBNA1 protein coding sequence is cloned into a baculovirus transfer vector. The baculovirus transfer vector and *Autographica californica* nuclear polyhedrosis genomic DNA are then co-transfected into insect cells, and recombinant baculoviruses are recovered. Cells are then infected with the recombinant baculovirus under conditions facilitating expression of isolated EBNA1 protein or polypeptide in the cell. In this process, the EBNA1 protein coding sequence includes no more than 90% of the Gly-Ala repeat amino acid sequence present in the naturally-occurring EBNA1 protein coding sequence which spans the Gly-Ala repeat amino acid sequence.

By utilizing this process, EBNA1 protein or polypeptide is expressed in quantities sufficient for the production of a detection immunoassay for EBV which provides few false positive readings.

## BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows the construction of the EBNA1 baculovirus transfer vector pVL941-EBNA1. The sequence of the oligonucleotide linkers inserted in the polyhedrin gene of the baculovirus transfer vector pVL941-SW is shown above the plasmid. The *underlined* ATG is the only ATG sequence in the 5' region of the polyhedrin gene and was used as the start codon for translation of the EBNA1 gene. After linearization of pVL941-SW with *Nco*I and digestion with bacterial alkaline phosphatase, the 3'-recessed ends were extended with the Klenow fragment of DNA polymerase I. The

EBNA1 gene was excised from p205 with *RsaI* and *BalI* enzymes, which remove the first seven codons of the gene, and ligated into pVL941-SW to form pVL941-EBNA1. Hygromycin phosphotransferase (*hph*) and $\beta$-lactamase (*amp*) genes are also shown.

FIG. 2 shows a modified protocol for improved yield and purity of bEBNA1. This is a Coomassie Blue stained SDS-polyacrylamide gel analysis of each step in the new purification scheme. The lanes read from right to left instead of from left to right. Starting at the far right the lane marked **"EBNA1"** is a lane of bEBNA1 protein purified by this procedure (i.e. it is the same as the lane on the far left) as verified by ability to bind to oriP. **Cells** - are whole SF-9 cells infected with the recombinant bEBNA1 recombinant baculovirus and the bEBNA1 band is visible. **Cytoplasm** - is the cytoplasmic supernatant after lysing the cells and spinning down the nuclei. **Nuclei** - is the whole nuclei after cell lysis and separating out nuclei from cytoplasm by centrifugation. **PolyminP** - is the supernatant after lysis of the nuclei and pelleting the DNA by PolyminP and centrifugation. **30% A.S.** - is the pellet that forms upon adding ammonium sulfate to the PolyminP supernatant (no significant bEBNA1 present). **45% A.S.** is the pellet that forms upon adding ammonium sulfate to the 30% A.S. supernatant to a final concentration of 45% (contains enriched bEBNA1). **60% A.S.** - is the pellet that forms upon adding ammonium sulfate to the 45% supernatant to a final concentration of 60% (no bEBNA1 present due to its being present only in the 30-45% cut). **Heparin** - bEBNA1 after chromotography of the 45% A.S. pellet over the Heparin Sepharose column. **Oligo. Aff.** - bEBNA1 after chromatography of the Heparin fraction over the oriP oligonucleotide affinity column.

FIG. 3 shows the phosphate labeling and phosphatase digestion of bEBNA1. Sf-9 cells were infected

with the AcMNPV-EBNA1 baculovirus and labeled with [$^{32}$P]orthophosphate as described in the Examples. Labeled cells were separated into cytoplasmic (*cyt*) and nuclear (*nuc*) fractions, and bEBNA1 was purified to homogeneity from the nuclear extract. Pure [$^{32}$P]EBNA1 was incubated at 25°C for 1 h either with (+) or without (-) CIP. Samples were subjected to electrophoresis on 12% SDS-polyacrylamide gels and $^{32}$P-labeled proteins were detected upon autoradiography of wet gels.

FIG. 4 shows the phosphoamino acid analysis of bEBNA1. Pure [32P]bEBNA1 was hydrolyzed in 6 N HCl (constant boiling) at 110°C for 1 or 2 h, then combined with unlabeled phosphoserine (*Ser-P*), phosphothreonine (*Thr-P*), and phosphotyrosine (*Tyr-P*) markers. The mixture of hydrolyzed amino acids and phosphoamino acid standards was separated by high voltage paper electrophoresis. Positions of phosphoamino acid markers were visualized by ninhydrin staining and are indicated by *dotted circles*. Positions of $^{32}$P-labeled amino acids and nonhydrolyzed bEBNA1 (0-h data lane) were identified by autoradiography.

FIGS. 5A and B show the native aggregation state of bEBNA1. bEBNA1 was combined with the protein standards apoferritin (*apo*; 440 kDa), IgG (158 kDa), bovine serum albumin (*BSA*; 66 kDa), ovalbumin (*ova*; 45 kDa) and myoglobin (*myo*; 17 kDa), then analyzed by glycerol gradient sedimentation (A) or gel filtration on Superose (B) as described herein. bEBNA1 was identified in column fractions by the nitrocellulose filter binding assay. The sedimentation coefficient (*s*) and Stokes radius of bEBNA1 were determined by comparison to the positions of protein standards of which the *s* values and Stokes radii are known.

FIG. 6 shows the stoichiometry of [$^{35}$S]bEBNA1 bound to *oriP* DNA. [$^{35}$S]bEBNA1 was incubated with pGEMoriP7, then gel-filtered to separate [$^{35}$S]bEBNA1 bound to pGEMoriP7 in the excluded fractions from unbound bEBNA1 in the included

fractions as described in the Examples. Fractions were analyzed for DNA and [$^{35}$S]bEBNA1.

FIG. 7 shows the salt dependence of bEBNA1 binding to the family of repeats and the dyad symmetry element. bEBNA1 (50 ng) was incubated with 40 fmol of $^{32}$P-end-labeled DNA containing either the dyad symmetry element (*closed circles*) or the family of repeats (*open circles*) in the presence of 2.5 $\mu$g of calf thymus DNA and various concentrations of NaCl. After 10 min at 23°C, the reaction mixture was filtered through nitrocellulose, and the DNA retained on the filters was quantitated by liquid scintillation.

FIG. 8 shows the effect of the family of repeats on binding of bEBNA1 to the dyad symmetry element. *Top*, diagram of *oriP* showing the disposition of EBNA1 binding sites (*boxes*). *Bottom*, 10 fmol of $^{32}$P-labeled DNA fragment containing either the family of repeats (*open circles*), the dyad symmetry element (*closed circles*), or the complete *oriP* (*closed triangles*) were incubated with various amounts of bEBNA1 (shown as fmol dimers) in 50 mM HEPES (pH 7.5), 300 mM NaCl, 5 mM $MgCl_2$ for 10 min at 23°C. Reactions containing the family of repeats or dyad symmetry element were then filtered through nitrocellulose. Reactions containing the complete *oriP* (*closed triangles*) were treated with 50 units of *Eco*RV for 3 min at 37°C to separate the family of repeats from the end-labeled dyad symmetry element (see scheme, *top*) prior to filtration through nitrocellulose.

FIGS. 9A and B show the protection of the *Ava*I site in the dyad symmetry element by bEBNA1. In FIG. 9A, the 300-bp DNA fragment containing the dyad symmetry element, $^{32}$P-end-labeled at one end only, was incubated with various amounts of bEBNA1 (shown as fmol dimers) prior to digestion with *Ava*I and electrophoresis on a 6% polyacrylamide gel. The DNA was visualized by autoradiography of dried gels. Scheme of DNA fragment (*top*)

shows EBNA1 consensus binding sites (*boxes*). In FIG. 9B, the *Ava*I-protected bands in the autoradiograph in *A* were quantitated by a laser densitometer (LKB Bromma Ultroscan XL).

FIG. 10 is cloning scheme for preparation of a vector for expression in *E. coli* of EBNA1.

FIG. 11 is a map for the plasmid p291. The *Hind*III fragment contains the eEBNA1 gene's nucleotides 107930-110493 (2.563kb) from the strain EBV B93-8, with the eEBNA1 gene itself spanning nucleotides 107950 to 109872 (1.922kb). The PCR product of the eEBNA1 gene between the N and C terminii PCR primers is shown in the upper right.

FIGS. 12A-C show the full double stranded DNA PCR product of the eEBNA1 gene with restriction endonuclease sites. The upper strand corresponds to SEQ. ID. No. 3.

## DETAILED DESCRIPTION

The present invention relates to a process for recovering EBNA1 protein or polypeptide having the following steps: providing cells having a nucleus containing EBNA1 protein or polypeptide; recovering the nucleus containing expressed EBNA1 protein or polypeptide from the cells; separating the nucleus containing expressed EBNA1 protein or polypeptide into a liquid fraction containing the expressed EBNA1 protein or polypeptide and a solid fraction containing substantially all DNA from the nucleus; separating the liquid fraction from the solid fraction; and recovering EBNA1 protein or polypeptide from the liquid fraction. In this process, the nucleus is separated by centrifugation where the liquid fraction is a supernatant and the solid fraction is a pellet. After centrifugation, the supernatant contains less than 5% of DNA.

The process further provides subjecting the liquid fraction to a first ammonium sulfate treatment at an

ammonium sulfate concentration which forms a solid phase containing contaminant proteins and a liquid phase containing EBNA1 protein or polypeptide, followed by subjecting the liquid phase containing EBNA1 protein or polypeptide to a second ammonium sulfate treatment at an ammonium sulfate concentration which forms a solid phase containing EBNA1 protein or polypeptide and a liquid phase containing contaminant proteins and then finally separating the solid phase containing EBNA1 protein or polypeptide and the liquid phase containing contaminant proteins. The first ammonium sulfate treatment is at a >0 to 30%, preferably 30%, ammonium sulfate concentration and the second ammonium sulfate treatment is at a 30 to 45%, preferably 45%, ammonium sulfate concentration.

The solid phase containing EBNA1 protein or polypeptide is then purified, after separation, by affinity column chromatography, such as agarose-heparin column chromatography or oligonucleotide affinity column chromatography. By utilizing this purification process, it is believed that the recovered EBNA1 protein is folded in its natural conformation.

This process produced abundant quantities of purified EBNA1 protein or polypeptide useful for diagnosis of EBV.

According to one embodiment, insect cells, preferably Sf-9 insect cells, are grown and infected with EBNA1-containing recombinant baculovirus, then harvested after a sufficient amount of time has passed to allow for protein expression. The cytoplasmic membrane is disrupted and the nuclei containing expressed baculovirus-derived EBNA1 protein or polypeptide ("bEBNA1") are pelleted to remove cytoplasm. The nuclei are lysed, producing a viscous solution ("nuclear extract") due to the presence of DNA. The DNA is then removed by sonication which shears the DNA and partially reduces the viscosity of the nuclear extract. A

chromatography preparation solution is then added to the nuclear extract which is incubated and then centrifuged. This packs the DNA down tight into a small pellet, leaving most of the solution free of DNA. The solution is decanted and then treated according to the above-described two-step ammonium sulfate precipitation procedure. The centrifugation procedure after the second ammonium sulfate precipitation step produced a supernatant which is discarded and a pellet with bEBNA1.

The pellet containing bEBNA1 is dissolved in a buffer and then dialyzed against the buffer. This dialyzed preparation is loaded onto an ion exchange chromatography column and eluted from it with a salt gradient and then purified using affinity column chromatography.

In another embodiment, *E. coli* cells, rather than insect cells, are used as host cells.

The present invention also relates to an isolated EBNA1 protein or polypeptide formulation having substantially no components which generate false positive readings when used to detect EBV in human serum. Furthermore, wherein naturally-occurring EBNA1 protein or polypeptide spans a Gly-Ala repeat amino acid sequence, the isolated EBNA1 protein or polypeptide of the present invention includes no more than 90%, preferably no more than 94%, of the Gly-Ala repeat amino acid sequence.

Additionally, the present invention provides an isolated DNA molecule encoding EBNA1 protein or polypeptide, a recombinant DNA expression system comprising an expression vector into which is inserted a heterologous DNA molecule encoding EBNA1 protein or polypeptide, and a host cell, such as an insect cell, incorporating a heterologous DNA molecule encoding EBNA1 protein or polypeptide, all of which have substantially no components which generate false positive readings when used to detect EBV in human serum.

The heterologous DNA molecule encoding the bEBNA1 protein or polypeptide of the present invention comprises the nucleotide sequence corresponding to SEQ. ID. No. 1 as follows:

ATG ACA GGA CCT GGA AAT GGC CTA GGA GAG
AAG GGA GAC ACA TCT GGA CCA GAA GGC TCC
GGC GGC AGT GGA CCT CAA AGA AGA GGG GGT
GAT AAC CAT GGA CGA GGA CGG GGA AGA GGA
CGA GGA CGA GGA GGC GGA AGA CCA GGA GCC
CCG GGC GGC TCA GGA TCA GGG CCA AGA CAT
AGA GAT GGT GTC CGG AGA CCC CAA AAA CGT
CCA AGT TGC ATT GGC TGC AAA GGG ACC CAC
GGT GGA ACA GGA GCA GGA GCA GGA GCG GGA
GGG GCA GGA GCA GGA GGT GGA GGC CGG GGT
CGA GGA GGT AGT GGA GGC CGG GGT CGA GGA
GGT AGT GGA GGC CGC CGG GGT AGA GGA CGT
GAA AGA GCC AGG GGG GGA AGT CGT GAA AGA
GCC AGG GGG AGA GGT CGT GGA CGT GGA GAA
AAG AGG CCC AGG AGT CCC AGT AGT CAG TCA
TCA TCA TCC GGG TCT CCA CCG CGC AGG CCC
CCT CCA GGT AGA AGG CCA TTT TTC CAC CCT
GTA GGG GAA GCC GAT TAT TTT GAA TAC CAC
CAA GAA GGT GGC CCA GAT GGT GAG CCT GAC
GTG CCC CCG GGA GCG ATA GAG CAG GGC CCC
GCA GAT CAC CCA GGA GAA GGC CCA AGC ACT
GGA CCC CGG GGT CAG GGT GAT GGA GGC AGG
CGC AAA AAA GGA GGG TGG TTT GGA AAG CAT
CGT GGT CAA GGA GGT TCC AAC CCG AAA TTT
GAG AAC ATT GCA GAA GGT TTA AGA GCT CTC
CTG GCT AGG AGT CAC GTA GAA AGG ACT ACC
GAC GAA GGA ACT TGG GTC GCC GGT GTG TTC
GTA TAT GGA GGT AGT AAG ACC TCC CTT TAC
AAC CTA AGG CGA GGA ACT GCC CTT GCT ATT
CCA CAA TGT CGT CTT ACA CCA TTG AGT CGT
CTC CCC TTT GGA ATG GCC CCT GGA CCC GGC

CCA CAA CCT GGC CCG CTA AGG GAG TCC ATT
GTC TGT TAT TTC ATG GTC TTT TTA CAA ACT
CAT ATA TTT GCT GAG GTT TTG AAG GAT GCG
ATT AAG GAC CTT GTT ATG ACA AAG CCC GCT
CCT ACC TGC AAT ATC AGG GTG ACT GTG TGC
AGC TTT GAC GAT GGA GTA GAT TTG CCT CCC
TGG TTT CCA CCT ATG GTG GAA GGG GCT GCC
GCG GAG GGT GAT GAC GGA GAT GAC GGA GAT
GAA GGA GGT GAT GGA GAT GAG GGT GAG GAA
GGG CAG GAG TGA

The amino acid sequence, corresponding to the DNA molecule of SEQ. ID. No. 1, is SEQ. ID. No. 2 as follows:

Met Thr Gly Pro Gly Asn Gly Leu Gly Glu
Lys Gly Asp Thr Ser Gly Pro Glu Gly Ser
Gly Gly Ser Gly Pro Gln Arg Arg Gly Gly
Asp Asn His Gly Arg Gly Arg Gly Arg Gly
Arg Gly Arg Gly Gly Gly Arg Pro Gly Ala
Pro Gly Gly Ser Gly Ser Gly Pro Arg His
Arg Asp Gly Val Arg Arg Pro Gln Lys Arg
Pro Ser Cys Ile Gly Cys Lys Gly Thr His
Gly Gly Thr Gly Ala Gly Ala Gly Ala Gly
Gly Ala Aly Ala Gly Gly Gly Gly Arg Gly
Arg Gly Gly Ser Gly Gly Arg Gly Arg Gly
Gly Ser Gly Gly Arg Arg Gly Arg Gly Arg
Glu Arg Ala Arg Gly Gly Ser Arg Glu Arg
Ala Arg Gly Arg Gly Arg Gly Arg Gly Glu
Lys Arg Pro Arg Ser Pro Ser Ser Gln Ser
Ser Ser Ser Gly Ser Pro Pro Arg Arg Pro
Pro Pro Gly Arg Arg Pro Phe Phe His Pro
Val Gly Glu Ala Asp Tyr Phe Glu Tyr His
Gln Glu Gly Gly Pro Asp Gly Glu Pro Asp
Val Pro Pro Gly Ala Ile Glu Gln Gly Pro
Ala Asp His Pro Gly Glu Gly Pro Ser Thr
Gly Pro Arg Gly Gln Gly Asp Gly Gly Arg
Arg Lys Lys Gly Gly Trp Phe Gly Lys His

Arg Gly Gln Gly Gly Ser Asn Pro Lys Phe
Glu Asn Ile Ala Glu Gly Leu Arg Ala Leu
Leu Ala Arg Ser His Val Glu Arg Thr Thr
Asp Glu Gly Thr Trp Val Ala Gly Val Phe
Val Tyr Gly Gly Ser Lys Thr Ser Leu Tyr
Asn Leu Arg Arg Gly Thr Ala Leu Ala Ile
Pro Gln Cys Arg Leu Thr Pro Leu Ser Arg
Leu Pro Phe Gly Met Ala Pro Gly Pro Gly
Pro Gln Pro Gly Pro Leu Arg Glu Ser Ile
Val Cys Tyr Phe Met Val Phe Leu Gln Thr
His Ile Phe Ala Glu Val Leu Lys Asp Ala
Ile Lys Asp Leu Val Met Thr Lys Pro Ala
Pro Thr Cys Asn Ile Arg Val Thr Val Cys
Ser Phe Asp Asp Gly Val Asp Leu Pro Pro
Trp Phe Pro Pro Met Val Glu Gly Ala Ala
Ala Glu Gly Asp Asp Gly Asp Asp Gly Asp
Glu Gly Gly Asp Gly Asp Glu Gly Glu Glu
Gly Gln Glu OPA

Production of this isolated protein or polypeptide is preferably carried out using recombinant DNA technology. Furthermore, the isolated DNA molecule is isolated from any other DNA molecule which expresses protein that generates false positive readings when the EBNA1 protein or polypeptide is used to detect EBV in human serum.

Additionally, the heterologous DNA molecule encoding the *E. coli* expression system-derived EBNA1 protein or polypeptide ("eEBNA1") of the present invention comprises the nucleotide sequence corresponding to SEQ. ID. No. 3 as follows:

ATG GGA GAA GGC CCA AGC ACT GGA CCC CGG
GGT CAG GGT GAT GGA GGC AGG CGC AAA AAA
GGA GGG TGG TTT GGA AAG CAT CGT GGT CAA
GGA GGT TCC AAC CCG AAA TTT GAG AAC ATT
GCA GAA GGT TTA AGA GCT CTC CTG GCT AGG
AGT CAC GTA GAA AGG ACT ACC GAC GAA GGA

ACT TGG GTC GCC GGT GTG TTC GTA TAT GGA
GGT AGT AAG ACC TCC CTT TAC AAC CTA AGG
CGA GGA ACT GCC CTT GCT ATT CCA CAA TGT
CGT CTT ACA CCA TTG AGT CGT CTC CCC TTT
GGA ATG GCC CCT GGA CCC GGC CCA CAA CCT
GGC CCG CTA AGG GAG TCC ATT GTC TGT TAT
TTC ATG GTC TTT TTA CAA ACT CAT ATA TTT
GCT GAG GTT TTG AAG GAT GCG ATT AAG GAC
CTT GTT ATG ACA AAG CCC GCT CCT ACC TGC
AAT ATC AGG GTG ACT GTG TGC AGC TTT GAC
GAT GGA GTA GAT TTG CCT CCC TGG TTT CCA
CCT ATG GTG GAA GGG GCT GCC GCG GAG GGT
GAT GAC GGA GAT GAC GGA GAT GAA GGA GGT
GAT GGA GAT GAG GGT GAG GAA GGG CAG GAG
CTG CGT CGT GCT TCT GTT GGT TAA

The amino acid sequence, corresponding to the DNA molecule of SEQ. ID. No. 3, is SEQ. ID. No. 4 as follows:

Met Gly Glu Gly Pro Ser Thr Gly Pro Arg
Gly Gln Gly Asp Gly Gly Arg Arg Lys Lys
Gly Gly Trp Phe Gly Lys His Arg Gly Gln
Gly Gly Ser Asn Pro Lys Phe Glu Asn Ile
Ala Glu Gly Leu Arg Ala Leu Leu Ala Arg
Ser His Val Glu Arg Thr Thr Asp Glu Gly
Thr Trp Val Ala Gly Val Phe Val Tyr Gly
Gly Ser Lys Thr Ser Leu Tyr Asn Leu Arg
Arg Gly Thr Ala Leu Ala Ile Pro Gln Cys
Arg Leu Thr Pro Leu Ser Arg Leu Pro Phe
Gly Met Ala Pro Gly Pro Gly Pro Gln Pro
Gly Pro Leu Arg Glu Ser Ile Val Cys Tyr
Phe Met Val Phe Leu Gln Thr His Ile Phe
Ala Glu Val Leu Lys Asp Ala Ile Lys Asp
Leu Val Met Thr Lys Pro Ala Pro Thr Cys
Asn Ile Arg Val Thr Val Cys Ser Phe Asp
Asp Gly Val Asp Leu Pro Pro Trp Phe Pro
Pro Met Val Glu Gly Ala Ala Ala Glu Gly

Asp Asp Gly Asp Asp Gly Asp Glu Gly Gly
Asp Gly Asp Glu Gly Glu Glu Gly Gln Glu
Leu Arg Arg Ala Ser Val Gly OCH

The DNA molecule encoding the EBNA1 protein or polypeptide of the present invention can be incorporated in cells using conventional recombinant DNA technology. Generally, this involves inserting the DNA molecule into an expression system to which the DNA molecule is heterologous (i.e. not normally present). The heterologous DNA molecule is inserted into the expression system or vector in proper orientation and correct reading frame. The vector contains the necessary elements for the transcription and translation of the inserted protein-coding sequences.

U.S. Patent No. 4,237,224 to Cohen and Boyer, which is hereby incorporated by reference, describes the production of expression systems in the form of recombinant plasmids using restriction enzyme cleavage and ligation with DNA ligase. These recombinant plasmids are then introduced by means of transformation and replicated in unicellular cultures including procaryotic organisms and eucaryotic cells grown in tissue culture.

Recombinant genes may also be introduced into viruses, such as vaccina virus. Recombinant viruses can be generated by transfection of plasmids into cells infected with virus.

Suitable vectors include, but are not limited to, the following viral vectors such as lambda vector system gt11, gt WES.tB, Charon 4, and plasmid vectors such as pBR322, pBR325, pACYC177, pACYC184, pUC8, pUC9, pUC18, pUC19, pLG339, pR290, pKC37, pKC101, SV 40, pBluescript II SK +/- or KS +/- (see "Stratagene Cloning Systems" Catalog (1993) from Stratagene, La Jolla, Calif, which is hereby incorporated by reference), pQE, pIH821, pGEX, pET series (see F.W. Studier et. al., "Use of T7 RNA Polymerase to Direct Expression of Cloned Genes," Gene Expression

Technology vol. 185 (1990), which is hereby incorporated by reference) and any derivatives thereof. Recombinant molecules can be introduced into cells via transformation, particularly transduction, conjugation, mobilization, or electroporation. The DNA sequences are cloned into the vector using standard cloning procedures in the art, as described by Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Springs Laboratory, Cold Springs Harbor, New York (1982), which is hereby incorporated by reference.

A variety of host-vector systems may be utilized to express the protein-encoding sequence(s). Primarily, the vector system must be compatible with the host cell used. Host-vector systems include but are not limited to the following: bacteria transformed with bacteriophage DNA, plasmid DNA, or cosmid DNA; microorganisms such as yeast containing yeast vectors; mammalian cell systems infected with virus (e.g., vaccinia virus, adenovirus, etc.); insect cell systems infected with virus (e.g., baculovirus). The expression elements of these vectors vary in their strength and specificities. Depending upon the host-vector system utilized, any one of a number of suitable transcription and translation elements can be used.

Different genetic signals and processing events control many levels of gene expression (e.g., DNA transcription and messenger RNA (mRNA) translation).

Transcription of DNA is dependent upon the presence of a promotor which is a DNA sequence that directs the binding of RNA polymerase and thereby promotes mRNA synthesis. The DNA sequences of eucaryotic promotors differ from those of procaryotic promotors. Furthermore, eucaryotic promotors and accompanying genetic signals may not be recognized in or may not function in a procaryotic system, and, further, procaryotic promotors are not recognized and do not function in eucaryotic cells.

Similarly, translation of mRNA in procaryotes depends upon the presence of the proper procaryotic signals which differ from those of eucaryotes. Efficient translation of mRNA in procaryotes requires a ribosome binding site called the Shine-Dalgarno (SD) sequence on the mRNA. This sequence is a short nucleotide sequence of mRNA that is located before the start codon, usually AUG, which encodes the amino-terminal methionine of the protein. The SD sequences are complementary to the 3'-end of the 16S rRNA (ribosomal RNA) and probably promote binding of mRNA to ribosomes by duplexing with the rRNA to allow correct positioning of the ribosome. For a review on maximizing gene expression, see Roberts and Lauer, Methods in Enzymology, 68:473 (1979), which is hereby incorporated by reference.

Promotors vary in their "strength" (i.e. their ability to promote transcription). For the purposes of expressing a cloned gene, it is desirable to use strong promotors in order to obtain a high level of transcription and, hence, expression of the gene. Depending upon the host cell system utilized, any one of a number of suitable promotors may be used. For instance, when cloning in *E. coli*, its bacteriophages, or plasmids, promotors such as the T7 phage promoter, *lac* promotor, *trp* promotor, *rec*A promotor, ribosomal RNA promotor, the $P_R$ and $P_L$ promotors of coliphage lambda and others, including but not limited, to *lac*UV5, *omp*F, *bla*, *lpp*, and the like, may be used to direct high levels of transcription of adjacent DNA segments. Additionally, a hybrid *trp-lac*UV5 (*tac*) promotor or other *E. coli* promotors produced by recombinant DNA or other synthetic DNA techniques may be used to provide for transcription of the inserted gene.

Bacterial host cell strains and expression vectors may be chosen which inhibit the action of the promotor unless specifically induced. In certain operons, the

addition of specific inducers is necessary for efficient transcription of the inserted DNA. For example, the *lac* operon is induced by the addition of lactose or IPTG (isopropylthio-beta-D-galactoside). A variety of other operons, such as *trp*, *pro*, etc., are under different controls.

Specific initiation signals are also required for efficient gene transcription and translation in procaryotic cells. These transcription and translation initiation signals may vary in "strength" as measured by the quantity of gene specific messenger RNA and protein synthesized, respectively. The DNA expression vector, which contains a promotor, may also contain any combination of various "strong" transcription and/or translation initiation signals. For instance, efficient translation in *E. coli* requires a Shine-Dalgarno (SD) sequence about 7-9 bases 5' to the initiation codon (ATG) to provide a ribosome binding site. Thus, any SD-ATG combination that can be utilized by host cell ribosomes may be employed. Such combinations include but are not limited to the SD-ATG combination from the *cro* gene or the *N* gene of coliphage lambda, or from the *E. coli* tryptophan E, D, C, B or A genes. Additionally, any SD-ATG combination produced by recombinant DNA or other techniques involving incorporation of synthetic nucleotides may be used.

Once the isolated DNA molecule has been cloned into an expression system, it is ready to be incorporated into a host cell. Such incorporation can be carried out by the various forms of transformation noted above, depending upon the vector/host cell system. Suitable host cells include, but are not limited to, bacteria, insect, virus, yeast, mammalian cells, and the like.

The present invention also provides a method of expressing an EBNA1 protein coding sequence in a cell. In this expression process, an EBNA1 protein coding sequence is

cloned into a baculovirus transfer vector. The baculovirus transfer vector and *Autographica californica* nuclear polyhedrosis genomic DNA are then co-transfected into insect cells, and recombinant baculoviruses are recovered. Cells are then infected with the recombinant baculovirus under conditions facilitating expression of isolated EBNA1 protein or polypeptide in the cell. In this process, the EBNA1 protein coding sequence includes no more than 90%, preferably no more than 94%, of the Gly-Ala repeat amino acid sequence present in the naturally-occurring EBNA1 protein coding sequence which spans the Gly-Ala repeat amino acid sequence.

The isolated EBNA1 protein or polypeptide formulation of the present invention can be utilized for detection of EBV in a sample of human tissue or body fluids. This detection process involves providing the isolated EBNA1 protein or polypeptide formulation as an antigen, contacting the sample with the antigen, and detecting any reaction which indicates EBV is present in the sample using an assay system. More specifically, this technique permits detection of EBV in a sample of the following tissue or body fluids: blood, spinal fluid, sputum, pleural fluids, urine, bronchial alveolor lavage, lymph nodes, bone marrow, or other biopsied materials.

In one embodiment, the assay system has a sandwich or competitive format. Examples of suitable assays include an enzyme-linked immunosorbent assay, a radioimmunoassay, a gel diffusion precipitin reaction assay, an immunodiffusion assay, an agglutination assay, a fluorescent immunoassay, a protein A immunoassay, or an immunoelectrophoresis assay.

## EXAMPLES

Example 1 - Cells and Virus

The wild-type baculovirus, *Autographica californica* nuclear polyhedrosis virus (AcMNPV), and *Spodoptera frugiperdi* (Sf-9) cells used to propogate the baculoviruses, were kindly provided by Dr. Ora Rosen (Sloan-Kettering Cancer Center), with permission from Dr. Max D. Summers (Texas A & M University). Sf-9 cells were grown as monolayer cultures in Grace's medium (Gibco Laboratories) with 0.33% yeastolate and 0.33% lactalbumin hydrolysate (Difco) supplemented with 10% fetal bovine serum.

Example 2 - Plasmids

pVL941-SW (see Figure 1)was constructed from pVL941 by Dr. Susan Wente in Dr. Ora Rosen's laboratory, by insertion of an *Nco*I/*Xba*I/*Spe*I linker into the *Bam*HI site of the polyhedrin gene in pVL941. As shown in plasmid p205, containing the EBNA1 gene with a 700 bp (± 20 bp) deletion in the Gly-Ala repeat region, was kindly provided by Dr. Bill Sugden. Plasmid pGEMoriP7 was constructed by ligating *Rsa*I/*Hin*dIII DNA linkers to the ends of the *Rsa*I fragment of p220.2 (kindly provided by Dr. Bill Sugden) containing *oriP* and the EBNA1 gene and inserting this DNA fragment into the *Hin*dIII site of pGEMII (Promega Biotec, Madison, WI). pGEMoriP was constructed from pGEMoriP7 using *Acc*I to excise 2 kilobase pairs of DNA containing the EBNA1 gene followed by religation to give pGEMoriP, which contains the entire *oriP* sequence.

Example 3 - Construction of the EBNA1 Recombinant Baculovirus (AcMNPV-EBNA1)

The EBNA1 gene was excised from p205 using *RsaI* and *BalI*, which remove the first seven codons. The initiating methionine was regenerated upon ligation into the baculovirus transfer vector pVL941-SW to yield pVL941/EBNA1 (Fig. 1). pVL941/EBNA1 and AcMNPV DNA were cotransfected into Sf-9 insect cells by the calcium phosphate precipitation method as described by Summers et al., Tex. Agric. Exp. Stn. Bull., 1555:27-31 (1987), which is hereby incorporated by reference. Five days post-transfection, serial dilutions of the medium from the transfected cells were plated with Sf-9 cells in 96-well plates. After amplification of the virus for 4 days, the cells were screened for the presence of virus containing the EBNA1 gene by dot blot analysis. The medium from a positive well was then used in a plaque assay according to Summers et al., Tex. Agric. Exp. Stn. Bull., 1555:27-31 (1987), which is hereby incorporated by reference, and a recombinant (non-occluded) plaque was picked, analyzed for the presence of the EBNA1 gene by dot blots, and subjected to one more round of plaque purification. Virus from one of the resulting recombinant plaques was amplified in Sf-9 cells. Total DNA was prepared from these cells, digested with restriction enzymes, and analyzed by Southern blot hybridizations to verify the presence of the complete EBNA1 *RsaI-BalI* fragment in the recombinant virus.

Example 4 - DNA Oligonucleotide Affinity Column

The oligonucleotide affinity column used in the procedure of Frappier, et al., "Overproduction, Purification, and Characterization of EBNA1, the Origin Binding Protein of Epstein-Barr Virus," The Journal of Biological Chemistry, 266(12):7819-26 (1991) was very

difficult to synthesize and to use. It had very low binding capacity and each prep. needed to be run over the column in several batches. The bEBNA1 that eluted was thus quite dilute and needed to be concentrated using either a heparin column or a MonoQ column. The following is an account of how to synthesize this improved column.

The oligonucleotide sequences were: OLIGO1 5'Biotin-GGGAAGCATATGCTACCC-3' (SEQ. I.D. No. 5); and OLIGO2 5'-GGGTAGCATATGCATATGCTTCCC-3' (SEQ. I.D. No. 6). 350 nmole of oligo1 and 440 nmole of oligo2 were mixed in 20ml of 10 mM Tris-HCl (pH 7.2), 0.3 M NaCl, and 0.03 M sodium citrate (final pH 8.5). The reaction was heated to 95°C for two minutes and allowed to cool to room temperature. The annealed oligonucleotide was incubated with 20 ml of a 1:1 slurry of strepavadin beads (Sigma Chemical Company) and rotated end over end for 12 hours at 4°C. The solution was then placed into a glass column, the beds were allowed to settle, followed by an extensive wash to remove unreacted oligonucleotide with 20 mM Hepes (pH 7.5), 0.5 mM EDTA, 10% glycerol, and 350 mM NaCl. This column had a capacity of approximately 0.7 mg of bEBNA-1 per ml of packed beads.

Example 5 - Nitrocellulose Filter Binding Assays

During purification, bEBNA1 was followed and quantitated by its ability to specifically retain a 900-bp fragment of *oriP* containing 20 copies of the 30-bp repeated sequence (family of repeats fragment) onto nitrocellulose filters. The family of repeats fragment was excised from pGEMoriP with *Eco*RI and *Nco*I, purified by agarose gel electrophoresis followed by electroelution, and quantitated by measuring the absorbance at 260 nm. The *oriP* repeat fragment was end-labeled by filling in the 3'-recessed ends using the Klenow fragment of DNA polymerase I with four dNTPs and [$\alpha$-$^{32}$P]TTP. Assays for bENBA1 were performed by

incubating an aliquot (20-200 ng of protein) of each fraction with 10-100 fmol of the end-labeled family of repeats fragment for 10 min at 23°C, in 25$\mu$l of 50 mM HEPES (pH 7.5), 5 mM $MgCl_2$, and 300 mM NaCl containing 2.5-5 $\mu$g of calf thymus DNA. Reaction mixtures were then diluted with 900 $\mu$l of 50 mM HEPES (pH 7.5), 5 mM $MgCl_2$ and immediately filtered through 0.45-$\mu$m HA filters (Millipore). The filters were dried and counted by liquid scintillation.

In nitrocellulose filter binding assays of bENBA1 with the dyad symmetry element of *oriP*, a 300-bp fragment of *oriP* containing the dyad and its four associated EBNA1 binding sites was incubated with bENBA1 as described for the family of repeats. This dyad symmetry element fragment was excised from pGEMoriP with *Hin*dIII and *Eco*RV, gel-purified, quantitated by absorbance at 260 nm, and end-labeled as described for the family of repeats fragment.

In assays in which bENBA1 was incubated with the complete *oriP* sequence, a 2-kilobase pair DNA fragment containing *oriP* was prepared from pGEMoriP and end-labeled near the dyad. This fragment was prepared by linearizing pGEMoriP with *Hin*dIII, filling in the 3'-recessed ends with [$\alpha$-$^{32}$P]TTP using the Klenow fragment of DNA polymerase I, then digesting with *Bam*HI. The *Hin*dIII to *Bam*HI fragment containing the complete *oriP* sequence was gel-purified and incubated with bEBNA1 as described for the family of repeats fragment.

Example 6 - Protein Determinations

Protein concentration was determined by the method of Bradford, Anal. Biochem., 72:248-254 (1976), which is hereby incorporated by reference, using bovine serum albumin as a standard. The concentration of pure bEBNA1 was determined by amino acid analysis (Sloan-Kettering Institute, Microchemistry Laboratory).

Example 7 - Phosphate and Methionine Labeling of bEBNA1

Sf-9 cells (2.8 x $10^8$ cells, 10 x 150-$cm^2$ flasks) were infected with recombinant EBNA1 baculovirus as described herein. Twenty-four hours post-infection, the media was replaced with phosphate-free or methionine-free Grace's media (Gibco) supplemented with 0.33% lactalbumin hydrolysate and 1 mCi of [$^{32}$P]orthophosphate or [$^{35}$S]methionine (Du Pont-New England Nuclear). Cells were labeled for 18 h before nuclei were prepared. Labeled bEBNA1 was purified as described herein.

Example 8 - Phosphoamino Acid Analysis of bEBNA1

Acid hydrolysis of bEBNA1 and resolution of phosphoamino acids was performed according to the method of Cooper et al., Methods Enzymol., 99:387-402 (1983), which is hereby incorporated by reference. Four-microgram samples of pure [$^{32}$P]bEBNA1 (4 $\mu$l) were added to 50 $\mu$l of 6 N constant boiling HCl (Pierce Chemical Co.) and heated to 110°C in a screw-cap 1.5-ml Eppendorf tube for 1, 2, or 4 h. The samples were lyophilized and resuspended in 2 $\mu$l of distilled water containing 4 $\mu$g each of phosphoserine, phosphothreonine, and phosphotyrosine markers. One microliter of each sample (2 $\mu$g of hydrolyzed bEBNA1) was spotted onto a strip of Whatman No. 3MM paper and subjected to electrophoresis in 0.5% pyridine, 5% acetic acid for 10 min at 2000 V. The paper was then dried and stained with ninhydrin to visualize the phosphoamino acid markers. $^{32}$P-Labeled amino acids of bEBNA1 were identified by autoradiography.

Example 9 - Phosphatase Treatment

Complete dephosphorylation of $^{32}P$-labeled bEBNA1 was achieved by treating 0.25 $\mu g$ of [$^{32}P$]bEBNA1 with 9 units of CIP (i.e., alkaline phosphatase from calf intestine) (Sigma) in 20 $\mu l$ of 10mM Tris-HCl (pH 8.0), 1 mM $MgCl_2$, 0.1 mM $ZnCl_2$, 300 mM NaCl for 1 h at 25°C.

Example 10 - Native Molecular Weight Determinations

The sedimentation coefficient of bENBA1 was measured by layering 40 $\mu g$ of bEBNA1 either alone or along with 60 $\mu g$ of molecular weight standards (apoferritin, IgG, bovine serum albumin, avalbumin, and myoglobin) in 200 $\mu l$ of 25 mM Tris-HCl (pH 7.5), 300 mM NaCl, 0.5 mM EDTA, 10% glycerol onto 12-ml 10-30% glycerol gradients containing 25 mM Tris-HCl (pH 7.5), 300 mM NaCl, 0.5 mM EDTA. Gradients were spun for 40 h at 270,000 x g at 5°C in a TH-641 rotor. After centrifugation, fractions of 160 $\mu l$ were collected from the bottom of each tube.

The Stokes radius of bENBA1 was determined by injecting 40 $\mu g$ of bEBNA1 along with 60 $\mu g$ of molecular weight standards in 200 $\mu l$ of 25 mM Tris-HCl (pH 7.5), 300 mM NaCl, 0.5 mM EDTA, 10% glycerol onto a 30-ml fast protein liquid chromatography Superose 12 gel filtration column. The column was developed in the same buffer. Fractions of 160 $\mu l$ were collected. Two microliters of each fraction from the glycerol gradients and gel filtration columns were assayed for the presence of bEBNA1 using the nitrocellulose filter binding assay described above. bENBA1 and the molecular weight standards were visualized after SDS-polyacrylamide gel electrophoresis analysis by staining with Coomassie Blue.

Example 11 - Stoichiometry of bEBNA1 on *oriP*

Thirty-five micrograms (7.7 pmol as plasmid circles) of pGEMoriP7 were incubated with excess $^{35}$S-labeled bEBNA1 (56 $\mu$g, 1.1 nmol as monomer) for 10 min at 37 °C in 200 $\mu$l of 20 mM HEPES (pH 7.5), 5 mM $MgCl_2$, 300 mM NaCl, 40% glycerol. The reaction was gel-filtered over a 5 ml Bio-Gel A-5m column at 4 °C in the same buffer, and 140-$\mu$l fractions were collected. [$^{35}$S]bEBNA1 in each fraction was quantitated by counting 30 $\mu$l in a scintillation counter. The molar quantity of DNA in each fraction was measured upon diluting 100 $\mu$l of column fraction with 400 $\mu$l of column buffer and measuring the absorbance at 260 nm (assuming 1 absorbance unit equals 50 $\mu$l/ml DNA). Approximately 90% of the radioactivity and absorbance at 260 nm was recovered after gel filtration.

Example 12 - *Ava*I Endonuclease Protection Assay

The 300-bp *Hin*dIII to *Eco*RV fragment of pGEMoriP containing the dyad symmetry element was end-labeled using the Klenow fragment of DNA polymerase I and [$\alpha$-$^{32}$P]TTP to fill in the *Hin*dIII end of the fragment. bEBNA1 was incubated with 10 fmol of the $^{32}$P-labeled dyad fragment in a 20-$\mu$l reaction containing 50 mM HEPES (pH 7.5), 300 mM NaCl, 5 mM $MgCl_2$ for 10 min at room temperature. The reactions were then diluted to 50 mM NaCl and incubated with 30 units of *Ava*I at 37 °C for 3 min. Digestions were stopped by the addition of SDS to 1%. Half of each reaction was then subjected to electrophoresis on a 6% polyacrylamide gel, which was dried prior to autoradiography.

Example 13 - Expression of EBNA1 in Baculovirus

The EBNA1 gene was excised from plasmid p205 and inserted into the pVL941-SW baculovirus transfer vector as described more fully above and as shown in Fig. 1. The resulting plasmid, pVL941-EBNA1, contained the EBNA1 gene, which translates into a 50 kDa protein lacking six amino-terminal amino acids and approximately 232 contiguous Gly-Ala residues of the Gly-Ala repeat region. Of these 232 amino acid residues, 6 were downstream of the Gly-Ala repeat such that there are still 13 of the 239 Gly-Ala residues remaining, representing 5.44%. Neither of these regions was essential for EBNA1-dependent replication *in vivo* when tested separately. Recombination of pVL941 with AcMNPV wild-type baculovirus DNA resulted in a recombinant baculovirus (AcMNPV-EBNA1) containing the EBNA1 gene controlled by the strong polyhedrin gene promoter. The EBNA1 protein or polypeptide produced by Ac-MNPV-EBNA1 is bENBA1. bEBNA1 is not a fusion protein, as the ENBA1 gene was placed directly adjacent to the only ATG sequence present in the 5' region of the polyhedrin gene in pVL941-SW (Fig. 1).

Initially, Sf-9 monolayers were infected with AcMNPV-EBNA1 and harvested at 24-h intervals to determine the time course of bEBNA1 expression. bEBNA1 protein levels peaked approximately 48 h post-infection as determined by the ability of whole cell extracts to specifically retain the *oriP* repeat fragment on a nitrocellulose filter. The level of *oriP* binding activity correlated with the appearance on Coomassie Blue-stained SDS-polyacrylamide gels of a 50 kDa protein that was not present in Sf-9 cells infected with wild-type baculovirus (data not shown).

Example 14 - Purification of bEBNA1 from Insect Cells

Sf-9 cells were seeded into 16 150-$cm^2$ culture flasks (3 x $10^7$ cells/flask) (Corning), allowed to attach, then infected with AcMNPV-EBNA1 at a multiplicity of infection of three. The cells were harvested 46 h post-infection, washed in 250 ml of ice-cold phosphate-buffered saline, and resuspended on ice in 70 ml of hypotonic buffer (20 mM HEPES (pH 7.5), 1 mM $MgCl_2$, 1 mM PMSf) using a Dounce homogenizer with pestle B. Nuclei were collected upon centrifugation at 1000 x g for 10 min at 5°C, washed in 70 ml of cold hypotonic buffer, and resuspended with the Dounce homogenizer and pestle B in 20 ml of 20 mM HEPES (pH 7.5), 1 M NaCl, 1% Nonidet P-40, 10% glycerol, 1 mM $MgCl_2$, 1 mM PMSF, followed by incubation for 1 h on ice. This nuclear extract is sonicated for 2 minutes to shear the DNA and partially reduce the viscosity. A solution of 5% Polymin P® (Polyethylenimine, average molecular weight 50,000, Sigma Chemical Co., St. Louis, Mo.) of molecular weight 5000 daltons was prepared in 20 mM Tris-HCl (pH 7.5) and then added to the nuclear extract to a final concentration of 0.25%, incubated for 30 minutes on ice and then spun for 30 minutes at 18,000 rpm in the SS-34 rotor (Sorvall). This packs the DNA down tight into a small pellet thereby leaving most of the solution approximately 95% free of DNA.

After removing the DNA, the supernatant is adjusted to 30%, saturation of ammonium sulfate (e.g., adding 8.6 volume of 100% saturated ammonium sulfate solution) to precipitate contaminant proteins but not bEBNA1. After slowly stirring for 1 hour at 4 °C, the preparation is spun at 15,000 rpm for 30 minutes at 4 °C. The supernatant is decanted and then ammonium sulfate is added to a final saturation of 45% (e.g., adding 7.8 ml of 100% saturated ammonium sulfate solution) in order to bring down the bEBNA1, yet leave other contaminants in solution.

After slowly stirring for 1 hour at 4 °C, the preparation is spun for 30 minutes at 15,000 rpm at 4 °C. After discarding the supernatant, the bEBNA1 protein-containing pellet is then dissolved in buffer A (20 mM Hepes (pH 7.5), 0.5 mM EDTA, 2 mM DTT, 1 mM PMSF, 20% glycerol) and dialyzed against 2 liters of buffer A for 4 hours at 4 °C and then against another 2 liters of buffer A overnight before loading onto a 30-ml heparin-agarose column (Bio-Rad).

All column chromatography procedures to follow were at 4°C. The heparin-agarose column was washed with 60 ml of buffer A containing 500 mM NaCl at 0.27 ml/min; then bEBNA1 was eluted with buffer A containing 1 M NaCl (see FIG. 2). Fractions of the 1 M NaCl eluate containing *oriP* binding activity were pooled (26 ml), diluted to 350 mM NaCl with buffer A, then loaded onto 9 ml of the DNA oligonucleotide affinity column. The DNA affinity column was washed with 18 ml of 350 mM NaCl, and then bEBNA1 was eluted using buffer A containing 2 M NaCl.

If necessary, to concentrate bEBNA1, the 2 M NaCl eluate containing 33% of the *oriP* binding activity (50 ml) was dialyzed against 500 mM NaCl, diluted with buffer A to a conductivity equivalent to 260 mM NaCl (105 ml), and loaded onto a 1-ml Mono Q column. bEBNA1 was eluted with buffer A containing 500 mM NaCl. Aliquots of active fractions (20 μl/tube) were stored at -70°C. Alternatively, the bEBNA1 can be concentrated by diluting the preparation with buffer A to a conductivity in the range of 250-300 mM NaCl and loaded onto a 1 ml Heparin Agarose column followed by elution using buffer A containing 1M NaCl.

| Fraction | Protein | Activity | Specific Activity | Purification | Yield |
|---|---|---|---|---|---|
| | mg | units | units/mg | -fold | % |
| Cytoplasm | 129 | | | | |
| Nuclear/Polyp | 64 | 31,050 | 470 | 1 | 100 |
| 0-30% AS | 23.9 | | | | |
| 30-45% AS | 16.6 | 22,500 | 1355 | 2.9 | 73 |
| Heparin | 11. | 18,100 | 1631 | 3.5 | 58 |
| oligo-affinity | 7.3 | 11,800 | 1616 | 3.4 | 38 |

This modified protocol gives about a 5-fold higher amount of the bEBNA-1 at the end of the procedure. The greater amount is probably due to recovery of more bEBNA1 from the nucleus due to the elimination of DNA using Polyamine P instead of high speed centrifugation. In effect, one obtains much more solution phase due to tight compaction of the DNA by Polyamine P. The purity at the end is undoubtedly better than in the previous protocol due to the ammonium sulfate cut, but it cannot be detected by specific activity, because the difference is only between 95% and 98% (or greater) purity. However, for use of this product in an ELISA assay, one never knows when a very small level of impurity will invalidate the assay. Thus, the more pure - the better - even if it is a difference in going from 98 to 99 percent.

Example 15 - Biochemical Assays of bEBNA1

Homogeneous bEBNA1 was assayed for the ability to hydrolyze ATP, GTP, CTP, UTP, dATP, dGTP, dCTP, and TTP in 1, 3, and 10 mM $MgCl_2$, in the absence of DNA and in the presence of either *oriP*-containing duplex DNA or single-stranded DNA. Nucleotide hydrolysis assays were performed by incubating 200 ng of bEBNA1 with 50 $\mu$M [$\alpha$-$^{32}$P]- or [$\gamma$-$^{32}$P] nucleoside triphosphate and deoxynucleoside triphosphate in 10 $\mu$l of 20 mM Tris-HCl (pH 7.5) and 1, 3, or 10 mM $MgCl_2$ for 30 min at 37°C. Additional assays for nucleoside

triphosphate and deoxynucleoside triphosphate hydrolysis were performed in the presence of 50 ng of bacteriophage M13 single-stranded DNA at the three $MgCl_2$ concentrations, as well as in the presence of 75 ng of pGEMoriP at the three $MgCl_2$ concentrations. ATPase activity was also tested in the presence of 2 and 8 mM sodium acetate. Samples (0.5 $\mu$l) of reaction mixtures were spotted on polyethyleneimine cellulose thin layer chromatography plates and developed in 0.8 M acetic acid, 0.8 M LiCl (when $\gamma$-$^{32}$P-labeling was used). Reaction products were identified by autoradiography. The $\tau$ subunit of *Escherichia coli* DNA polymerase III holoenzyme was used as a positive control for ATP hydrolysis according to the method of Tsuchihashi et al., J. Biol. Chem., 264:17790-17795 (1989), which is hereby incorporated by reference. No hydrolysis of any nucleoside triphosphate by bEBNA1 was detected (data not shown).

Although all known helicases are ATPases, bEBNA1 was tested in the standard oligonucleotide displacement type of helicase assay according to Matson, J. Biol. Chem., 261:10169-10175 (1986), which is hereby incorporated by reference. bEBNA1 was examined for an ability to displace, from single-stranded circular bacteriophage $\phi$X174 DNA, a $^{32}$P-end-labeled flush DNA 30-mer, a 5'-tailed DNA 30-mer, and a 3'-tailed DNA 46-mer. In separate experiments three different synthetic DNA oligonucleotides were hybridized to bacteriophage $\phi$X174 single-stranded DNA to give either 1) flush (30-mer), 2) 5'-tailed (30-mer with 20 nucleotides annealed), or 3) 3'-tailed (46-mer with 30 nucleotides annealed) helicase substrates. The annealed oligonucleotides were 3'-end-labeled using either [$\alpha$-$^{32}$P]dCTP and the Klenow fragment of DNA polymerase I (flush and 5'-tailed substrates) or using terminal transferase (3'-tailed substrate). Each helicase substrate was then purified from unhybridized oligonucleotide by gel filtration on Bio-Gel A-1.5m. Helicase assays were performed by incubating 400 ng of

bEBNA1 with 9 fmol of DNA substrate in 30 mM HEPES (pH 7.5), 4 mM ATP, 7 mM $MgCl_2$, 1 mM dithiothrietol for 30 min at 37°C. Positive control reactions contained 400 ng of SV40 large T antigen. Reaction products were analyzed for oligonucleotide displacement on a 15% polyacrylamide gel. The SV40 large T antigen was used as a positive control according to the method of Goetz et al., J. Biol. Chem., 263:383-392 (1988), which is hereby incorporated by reference. Although the SV40 T antigen displaced each of these DNA oligonucleotides, no helicase activity was detected for baculoEBNA1, consistent with its lack of ATPase activity. Also tested were bEBNA1 for DNA polymerase, DNA ligase, endonuclease, exonuclease, and topoisomerase activities without positive results (not shown).

Note that in the following Examples 16-20 the bEBNA1 protein was purified by the process disclosed in L. Frappier, et.al., "Overproduction, Purification, and Characterization of EBNA1, the Origin Binding Protein of Espstein-Barr Virus," J. Biol. Chem. 766(12):7819-26 (1991) rather than by the process of Example 14.

Example 16 - Phosphoamino Acid Analysis of bEBNA1

bEBNA1 was labeled *in vivo* with [$^{32}$P]orthophosphate and purified to homogeneity. bEBNA1 was the major $^{32}$P-labeled protein in the nuclear extract and was not detected in the cytoplasm (Fig. 3). Treatment of pure [$^{32}$P]bEBNA1 with CIP resulted in loss of all detectable radioactive phosphate from bEBNA1 (Fig. 3). Since CIP has previously been shown to dephosphorylate serine residues only, Shaw et al., Virology, 115:88-96 (1981) and Klausing et al, Virol., 62:1258-1265 (1988), which are hereby incorporated by reference, bEBNA1 is presumably phosphorylated only on serine.

Further identification of phosphorylated residues in bEBNA1 was performed by acid hydrolysis of [$^{32}$P]bEBNA1 and separation of the phosphoamino acids by high voltage paper electrophoresis (Fig. 4). Samples of [$^{32}$P]bEBNA1 hydrolyzed for 1, 2, and 4 h were analyzed to ensure identification of any [$^{32}$P]phosphothreonine, which requires longer hydrolysis times, or [$^{32}$P]phosphotyrosine, which is less stable to acid hydrolysis according to the method of Cooper et al., Methods Enzymol., 99:387-402 (1983), which is hereby incorporated by reference. Upon electrophoresis, all the radioactive phosphate in bEBNA1 migrated in the position of phosphoserine; no radioactivity was detected at positions of phosphothreonine or phosphotyrosine (Fig. 4). After 4 h of acid hydrolysis, most of the radioactive phosphate was detected as free phosphate (data not shown).

Example 17 - Native Molecular Mass

bEBNA1 was analyzed by glycerol gradient sedimentation; an *s* value of 4.6 was obtained by comparison with protein markers with known *s* values (Fig. 3A). A Stokes radius of 50 Å for bENBA1 was determined by gel filtration analysis and comparison with protein standards of known Stokes radius (Fig. 3B). In both glycerol gradient and gel filtration analyses, *oriP* binding activity co-eluted with the bEBNA1 protein visualized in SDS-polyacrylamide gel analysis of the column fractions (data not shown). The *s* value and Stokes radius were combined in the equation of Siegel et al., Biochim. Biophys. Acta, 112:346-362 (1966), which is hereby incorporated by reference, to calculate a native molecular mass of 94 kDa for bEBNA1. The amino acid sequence of EBNA1 deduced from the DNA sequence of the EBNA1 gene predicts a molecular mass of about 41,309 kDa for a bEBNA1 monomer. Hence, the native molecular mass of bEBNA1 indicates that bEBNA1 is a dimer.

Example 18 - Stoichiometry of bEBNA1 Binding to *oriP*

$^{35}S$-Labeled bEBNA1 protein was prepared *in vivo* by metabolic labeling using [$^{35}S$]methionine followed by purification to homogeneity. The [$^{35}S$]bEBNA1 was used to measure the number of bEBNA1 molecules bound to *oriP* under conditions of saturating bEBNA1. A plasmid containing the complete *oriP* sequence was incubated with increasing amounts of [$^{35}S$]bEBNA1 then gel-filtered to separate [$^{35}S$]bEBNA1 bound to DNA in the excluded fractions from the unbound [$^{35}S$]bEBNA1 in the included fractions. Upon saturation of *oriP* with bEBNA1, indicated by the appearance of bEBNA1 monomers per *oriP* DNA which comigrated in the excluded fractions was 56 to 1 (Fig. 6). Since there are 24 EBNA1 binding sites in *oriP*, the stoichiometry of 2.3 bEBNA1 monomers per EBNA1 binding site indicates that bEBNA1 bound its site as a dimer, consistent with the native molecular weight of bEBNA1 and the palindromic structure of the consensus EBNA1 binding site.

Example 19 - Effect of Salt on Binding of bEBNA1 to the Family of Repeats and Dyad Symmetry Element

The effect of NaCl concentration on bEBNA1 binding to the family of repeats or the dyad symmetry element of *oriP* was studied using the nitrocellulose filter binding assay. bEBNA1 (50 ng) was incubated with 40 fmol of $^{32}P$-labeled dyad fragment or $^{32}P$-labeled repeat fragment in various concentrations of NaCl and in the presence of excess (2.5 $\mu$g) calf thymus DNA (Fig. 7). The binding profile indicates that the specific interaction of bEBNA1 with the dyad symmetry element was maximum at 250-300 mM NaCl and dropped off sharply at higher NaCl concentrations. Binding of bEBNA1 to the family of repeats, however, remained stable up to 500 mM NaCl. Hence, the relative binding strength of

bEBNA1 for the family of repeats *versus* the dyad symmetry element depended on the salt concentration. The apparent requirement of high salt for binding bEBNA1 to labeled DNA in these experiments may be attributed to efficient competition by nonspecific calf thymus DNA at low NaCl concentration.

Example 20 - bEBNA1 Binding to the Dyad Symmetry Element

The interaction of bEBNA1 with the family of repeats and dyad symmetry element of *oriP* was also assessed by examining the amount of bEBNA1 required to retain each element on nitrocellulose filters. Increasing amounts of bEBNA1 were incubated with 10 fmol of $^{32}$P-end-labeled repeat or dyad DNA fragment in 20 $\mu$l of buffer containing 300 mM NaCl and no calf thymus DNA. Retention of the dyad symmetry element onto nitrocellulose appeared to have a threshold where significant retention was not observed below 20 bEBNA1 dimers per dyad fragment (200 ng in Fig. 8, *closed circles*), but full retention was achieved at 50 bEBNA1 dimers per dyad (500 ng in Fig. 8). It would seem from this behavior that bEBNA1 must reach a critical concentration before it binds the dyad symmetry element. The apparent $K_d$ for bEBNA1 binding to the dyad symmetry element calculated from these data is 2 nM (assuming four bEBNA1 dimers were bound per dyad symmetry element). The family of repeats was retained onto nitrocellulose at lower levels of bEBNA1 than required for binding the dyad symmetry element (Fig. 8, *open circles*). An apparent $K_d$ for bEBNA1 binding to the family of repeats was calculated to be 0.2 nM (assuming four bEBNA1 dimers were bound per family of repeats).

The binding of bEBNA1 to the dyad symmetry element was further examined by an *AvaI* endonuclease protection assay. An *AvaI* site was present at the junction of two of the four EBNA1 binding sites in the dyad symmetry element

(Fig. 9). Increasing amounts of bEBNA1 were incubated with 10 fmol of the dyad symmetry element, end-labeled with $^{32}P$ at one end only. The reaction was then treated with sufficient *Ava*I to completely digest the DNA within 3 min at 37°C. Digestions were stopped with SDS and subjected to polyacrylamide gel electrophoresis to separate DNA fragments cut by *Ava*I from uncut (*Ava*I-protected) DNA (Fig. 9). As in the nitrocellulose binding assay, the *Ava*I protection analysis showed that a 20-fold molar excess of bEBNA1 dimers (200 ng in Fig. 9) was required over the dyad fragment to detect protection of the *Ava*I site, followed by a very sharp increase in protection against *Ava*I at levels above 20 bEBNA1 dimers per dyad symmetry element. The small difference between the *Ava*I protection assay (Fig. 9) and the nitrocellulose filter binding assay (Fig. 8) showed approximately 1.5 times more bEBNA1 was needed to bind the dyad symmetry element onto a nitrocellulose filter relative to the amount of bEBNA1 needed to protect the *Ava*I site. This may be due to the requirement for bEBNA1 to bind to only one particular site in the dyad symmetry element to protect it from *Ava*I, whereas retention of the dyad onto nitrocellulose may require bEBNA1 bound to another site or multiple bEBNA1 molecules bound to multiple sites in the dyad symmetry element.

*In vivo* the dyad symmetry element is accompanied by the family of repeats within *oriP* which may affect the interaction of EBNA1 with the dyad symmetry element in the complete *oriP* sequence. bEBNA1 was incubated with *oriP* labeled with $^{32}P$ at the end near the dyad. Just prior to filtration through nitrocellulose, the family of repeats was separated from the dyad symmetry element by digestion with *Eco*RV (Fig. 8) for each assay an aliquot was removed prior to filtration, quenched with SDS (i.e., sodium dodecyl sulfate), and analyzed in an agarose gel to confirm that *Eco*RV had completely separated the dyad from the *oriP* DNA.

The results showed significant amounts of dyad symmetry element were retained onto nitrocellulose at lower levels of bEBNA1 (200 fmol and less) in the presence of the family of repeats than in their absence (Fig. 8, *closed triangles*). However, further along in the titration, more bEBNA1 was required to bind the dyad on *oriP* than to bind the dyad alone. Complete retention onto nitrocellulose of the isolated family of repeats and dyad symmetry fragments required 300 and 500 fmol of bEBNA1, respectively. Hence, it seems a paradox that even 800 fmol of bEBNA1 was not sufficient to retain onto nitrocellulose more than half of the dyad fragment when it was within the context of *oriP*. Possible explanations include the following. The presence of the family of repeats may destabilize the interaction of bEBNA1 with the dyad. A less stable complex of bEBNA1 with the dyad may assemble in the presence of the family of repeats. The nonessential region of *oriP* between the family of repeats and dyad symmetry element may influence the nitrocellulose binding assay, or the presence of the dyad may cause more cooperative binding of bEBNA1 to the family of repeats, effectively decreasing the availability of bEBNA1 for binding the dyad.

The above examples describe the overproduction of EBNA1, the viral encoded protein which binds the latent phase origin (*oriP*) of EBV, in the baculovirus system and its purification of homogeneity. Like EBNA1 from latently infected B cell lines (see Jones et al., J. Virol., 63:101-110 (1989); Hearing et al., Virology, 145:105-116 (1985); and Gahn et al., Cell, 58:527-535 (1989), which are hereby incorporated by reference) the bEBNA1 bound tightly to *oriP*, arrested replication forks within or near the *oriP* family of repeats and was phosphorylated on serine residues. Since phosphorylation can modulate protein function (see Donaldson et al., Proc. Natl. Acad. Sci. U.S.A., 84:759-763 (1987); Gould et al., Nature, 342:39-45 (1989); and McVey et al.,

Nature, 341:503-507 (1989), which are hereby incoporated by reference), it seems likely that initiation of replication from *oriP* will be regulated by phosphorylation of EBNA1.

The palindromic nature of each EBNA1 consensus site suggests that EBNA1 binds its DNA site as a dimer. Indeed bEBNA1 appeared to be a dimer in solution and the stoichiometry of 56 bEBNA1 molecules per 24 EBNA1 binding sites in the *oriP* sequence was consistent with EBNA1 binding its site as a dimer as predicted. See Ambinder et al., J. Virol., 64:2369-2379 (1990), which is hereby incorporated by reference.

Increasing evidence suggests replication initiates within the dyad symmetry element of *oriP*. See Gahn et al., Cell, 58:527-535 (1989) and Wysokinski et al., J. Virol., 63:2657-2666 (1989), which are hereby incorporated by reference. Replication initiation in the dyad is greatly stimulated by the family of repeats. Id. One mechanism by which the repeats might activate the dyad is by altering the interaction of EBNA1 with the dyad symmetry element. The nitrocellulose filter binding assay suggested that the family of repeats reduced the concentration of bEBNA1 required to initiate binding to the dyad of bEBNA1 required to initiate binding to the dyad symmetry element. If the interaction of EBNA1 with the dyad symmetry element is important for the initiation of replication from *oriP*, then the stimulation of dyad binding by the family of repeats at low EBNA1 concentration may be one mechanism by which the repeats enhance replication from *oriP*.

EBNA1 is essential for latent EBV replication, yet the precise biochemcal function of EBNA1 remains elusive. The bEBNA1 protein should prove useful in biochemical assays to analyze the mechanism by which EBNA1 activates *oriP* to function as an origin of replication, a plasmid maintenance element, and a transcriptional enhancer. See Yates et al., Cancer Cells, 6:197-205 (1988), which is hereby incorporated

by reference. Applicant finds no ATPase (or other nucleoside triphosphatase), helicase, ligase, topoisomerase, DNA polymerase, oxonuclease, or endonuclease activities associated with bEBNA1. The absence of ATPase and helicase activity suggests EBNA1 plays a different role in replication than the large T antigen of SV40. It is always possible, however, that the true activity of EBNA1 will only be revealed upon binding other proteins or by modification at a specific site(s). Furthermore, the possibility cannot be excluded that, although the six amino-terminal amino acids and glycine-alanine repeat region of EBNA1, lacking in bEBNA1, are nonessential for EBNA1 function *in vivo*, id., they may affect the biochemical activity of EBNA1 *in vitro*.

Elucidation of the precise role of EBNA1 in replication and the mechanism(s) of replication control at *oriP* would be greatly facilitated by development of an *in vitro* system capable of initiating replication from *oriP*.

Example 21 - Determination of Nucleotide and Amino Acid Sequences

The DNA template used for the sequence analysis of the GlyAla deletion was the 10.6 kb bEBNA1 baculovirus transfer vector, called pVL941-EBNA1, the construction of which was described in L. Frappier, et al., "Overproduction, Purification, and Characterization of EBNA1, the Origin Binding Protein of Epstein-Barr Virus," J. Biol. Chem. 766(12):7819-26 (1991). The sequencing primer used in this analysis was positioned 187 nucleotides in from the A of the ATG start codon of EBNA1; the sequence of the sequencing primer was 5'AAAAACGTCCAAGTTGCATTG-3' (SEQ. ID. No. 7). Sequencing was performed using the Sequenase based protocol and version 2 kit of United States Biochemical, Cleveland, Ohio according to the manufacturers specifications.

Example 22 - Expression and Purification of eEBNA1

The gene and expression plasmid were constructed by PCR using the following primers: N - terminus - 5′ - GAT CGG CAT ATG GGA GAA GGC CCA AGC ACT GGA - 3′ (the underline is the Met for the first amino acid, and the GGA that follows encodes amino acid 442 of EBNA1) (SEQ. ID. No. 8); and C - terminus - 5′ - CT GGT GGA TCC TTA ACC AAC AGA AGC ACG ACG CAG CTC CTG CCC TTC CTC AC - 3′ (the underlined codon encodes the last amino acid of the eEBNA1) (SEQ. ID. No. 9).

The template used in the PCR reaction was p291 (FIG. 11), a plasmid containing the entire EBNA1 gene (see FIGS. 12A-C). The cycling conditions were 94 °C, 30 sec./ 60 °C, 30 sec./ 72 °C, 60 sec. This cycle is repeated 30 times in 100 μliters of 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 200 μmolar each dATP, dCTP, dGTP, dTTP, 0.01% gelatin, 2.5 units *Taq*I polymerase (Perkin-Elmer Cetus), 1 μmolar of each primer (described above), and 1 ng of plasmid p291. After the PCR reaction, the 641 bp fragment was purified by phenol extraction in 2% SDS followed by sequential digestion with 10 units of NdeI (New England Biolabs) and then 10 units of *Bam*HI (New England Biolabs). The NdeI/*Bam*HI 624 bp fragment (see SEQ. ID. NO. 3) was purified from an agarose gel and ligated into pET3c (digested with NdeI and *Bam*HI) to yield pET-eEBNA1, as shown in Figure 10. Sequence analysis confirmed that no errors had been introduced by PCR amplification.

To express eEBNA1, the pET-eEBNA1 plasmid was transformed into *E. coli* strain BL21(DE3)pLysS and the cells were grown at 37 °C in 4 liters of LB medium (per liter: 10g Bacto-tryptone, 5g Bacto-yeast, 10g NaCl, pH 7.5) supplemented with 1% glucose, 10 μg/ml thiamine, 50 μg/ml thymine, 100 μg/ml ampicillin, and 30 μg/ml chloramphenicol. Upon reaching an absorbance at 600 nm of 0.8, IPTG was added

to 0.4 mM, and after 2 hours at 37 °C, the cells were harvested by centrifugation (15g net weight).

The cells were frozen at -70 °C and then thawed to 4 °C, and then resuspended in 40 ml of 25 mM Tris-HCl (pH 8.0), 10 mM EDTA, and 50 mM glucose. At this point, the cells lyse due to the lysozyme produced by the pLysS plasmid and the freeze-thaw procedure. The volume was brought to 100 ml using solution I and the DNA removed by precipitation by adding 10 ml NaCl, 1.4 ml of 5% Polymin P® (50 kDa) dissolved in 20 mM Tris-HCl (pH 7.5). After stirring slowly for 30 minutes at 4 °C, the precipitation was spun at 18,000 rpm at 4 °C.

The supernatant (82 ml) was adjusted to 70% ammonium sulfate by adding 191 ml of 100% saturated ammonium sulfate to precipitate the eEBNA1 protein. The eEBNA1-containing precipitate was then pelleted by centrifugation for 30 minutes at 1,000 rpm in the GSA rotor at 4 °C. The pellet was dissolved in 40 ml of buffer B (20 mM Tris-HCl (pH 7.5), 0.5 mM EDTA, 2 mM DTT, 20% glycerol, 0.1 mM phenylmethylsulfonyl fluoride (i.e., PMSF)) and loaded onto a 330 ml column of Bio-Gel P-6 equilibrated in buffer B. Fractions of 8 ml were collected at a flow rate of 3 ml/minute and assayed for total protein by the Bradford reagent (Bio-Rad). Peak fractions (11-29) are pooled (700 mg protein).

The 700 mg protein pool was loaded onto a 320 ml column of Heparin-Agarose (Bio-Rad) equilibrated in buffer B. The column was eluted with a 3.2 liter linear gradient of buffer B from 0 mM NaCl to 800 mM NaCl. Fractions of 26 ml were collected and assayed for total protein and for eEBNA1. The eEBNA1 eluted in fractions 60-96 and these were pooled (39 mg) and precipitated by adding 434g solid ammonium sulfate (70% saturation). The protein precipitate was collected by centrifugation, resuspended in 20 ml buffer B, and dialyzed against 2 liters of buffer B for 4 hours and

then against another 2 liters of buffer B overnight. The dialysate was loaded onto a 40 ml column of Q Sepharose (Pharmacia) equilibrated in buffer B. The eEBNA1 was eluted with a linear gradient of 400 ml of 0 mM NaCl to 800 ml NaCl in buffer B. Fractions of 5 ml were collected at a flow rate of 1 ml/minute and the fractions were assayed for eEBNA1. Fractions containing eEBNA1 were pooled (fractions 34-44, 20 mg total). This eEBNA1-containing pool had a conductivity equal to 386 mM NaCl and was diluted with buffer B to a conductivity equal to 48 mM NaCl, then loaded onto a 4 ml column of CM Sepharose (Pharmacia) equilibrated in buffer B. The eEBNA1 was eluted using a 40 ml linear gradient of 0 mM NaCl to 700 mM NaCl in buffer B and the fractions containing eEBNA1 were pooled (fractions 24-34, 18 mg total) and dialyzed against buffer B and stored frozen at -70 °C.

Although the invention has been described in detail for the purpose of illustration, it is understood that such detail is solely for that purpose, and variations can be made therein by those skilled in the art without departing from the spirit and scope of the invention which is defined by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Cornell Research Foundation, Inc.

(ii) TITLE OF INVENTION: Epstein-Barr Virus Nuclear Antigen 1 Protein and Its Expression and Recovery

(iii) NUMBER OF SEQUENCES: 9

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Nixon, Hargrave, Devans & Doyle
(B) STREET: Clinton Square, P.O. Box 1051
(C) CITY: Rochester
(D) STATE: New York
(E) COUNTRY: USA
(F) ZIP: 14603

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Goldman Esq., Michael L.
(B) REGISTRATION NUMBER: 30,727
(C) REFERENCE/DOCKET NUMBER: 19603/271 (D-1530)

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (716) 263-1304
(B) TELEFAX: (716) 263-1600

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1212 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATGACAGGAC CTGGAAATGG CCTAGGAGAG AAGGGAGACA CATCTGGACC AGAAGGCTCC      60
GGCGGCAGTG GACCTCAAAG AAGAGGGGGT GATAACCATG GACGAGGACG GGGAAGAGGA     120
CGAGGACGAG GAGGCGGAAG ACCAGGAGCC CCGGGCGGCT CAGGATCAGG GCCAAGACAT     180
AGAGATGGTG TCCGGAGACC CCAAAAACGT CCAAGTTGCA TTGGCTGCAA AGGGACCCAC     240
GGTGGAACAG GAGCAGGAGC AGGAGCGGGA GGGGCAGGAG CAGGAGGTGG AGGCCGGGGT     300
CGAGGAGGTA GTGGAGGCCG GGGTCGAGGA GGTAGTGGAG GCCGCCGGGG TAGAGGACGT     360
GAAAGAGCCA GGGGGGGAAG TCGTGAAAGA GCCAGGGGGA GAGGTCGTGG ACGTGGAGAA     420
AAGAGGCCCA GGAGTCCCAG TAGTCAGTCA TCATCATCCG GGTCTCCACC GCGCAGGCCC     480
CCTCCAGGTA GAAGGCCATT TTTCCACCCT GTAGGGGAAG CCGATTATTT TGAATACCAC     540
CAAGAAGGTG GCCCAGATGG TGAGCCTGAC GTGCCCCCGG GAGCGATAGA GCAGGGCCCC     600
GCAGATCACC CAGGAGAAGG CCCAAGCACT GGACCCCGGG GTCAGGGTGA TGGAGGCAGG     660
CGCAAAAAAG GAGGGTGGTT TGGAAAGCAT CGTGGTCAAG GAGGTTCCAA CCCGAAATTT     720
GAGAACATTG CAGAAGGTTT AAGAGCTCTC CTGGCTAGGA GTCACGTAGA AAGGACTACC     780
GACGAAGGAA CTTGGGTCGC CGGTGTGTTC GTATATGGAG GTAGTAAGAC CTCCCTTTAC     840
AACCTAAGGC GAGGAACTGC CCTTGCTATT CCACAATGTC GTCTTACACC ATTGAGTCGT     900
CTCCCCTTTG GAATGGCCCC TGGACCCGGC CCACAACCTG GCCCGCTAAG GGAGTCCATT     960
GTCTGTTATT TCATGGTCTT TTTACAAACT CATATATTTG CTGAGGTTTT GAAGGATGCG    1020
ATTAAGGACC TTGTTATGAC AAAGCCCGCT CCTACCTGCA ATATCAGGGT GACTGTGTGC    1080
AGCTTTGACG ATGGAGTAGA TTTGCCTCCC TGGTTTCCAC CTATGGTGGA AGGGGCTGCC    1140
GCGGAGGGTG ATGACGGAGA TGACGGAGAT GAAGGAGGTG ATGGAGATGA GGGTGAGGAA    1200
GGGCAGGAGT GA                                                        1212
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 404 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Thr Gly Pro Gly Asn Gly Leu Gly Glu Lys Gly Asp Thr Ser Gly
1 5 10 15

Pro Glu Gly Ser Gly Gly Ser Gly Pro Gln Arg Arg Gly Gly Asp Asn
20 25 30

His Gly Arg Gly Arg Gly Arg Gly Arg Gly Arg Gly Gly Gly Arg Pro
35 40 45

Gly Ala Pro Gly Gly Ser Gly Ser Gly Pro Arg His Arg Asp Gly Val
50 55 60

Arg Arg Pro Gln Lys Arg Pro Ser Cys Ile Gly Cys Lys Gly Thr His
65 70 75 80

Gly Gly Thr Gly Ala Gly Ala Gly Ala Gly Gly Ala Xaa Ala Gly Gly
85 90 95

Gly Gly Arg Gly Arg Gly Gly Ser Gly Gly Arg Gly Arg Gly Gly Ser
100 105 110

Gly Gly Arg Arg Gly Arg Gly Arg Glu Arg Ala Arg Gly Gly Ser Arg
115 120 125

Glu Arg Ala Arg Gly Arg Gly Arg Gly Arg Gly Glu Lys Arg Pro Arg
130 135 140

Ser Pro Ser Ser Gln Ser Ser Ser Ser Gly Ser Pro Pro Arg Arg Pro
145 150 155 160

Pro Pro Gly Arg Arg Pro Phe Phe His Pro Val Gly Glu Ala Asp Tyr
165 170 175

Phe Glu Tyr His Gln Glu Gly Gly Pro Asp Gly Glu Pro Asp Val Pro
180 185 190

Pro Gly Ala Ile Glu Gln Gly Pro Ala Asp His Pro Gly Glu Gly Pro
195 200 205

Ser Thr Gly Pro Arg Gly Gln Gly Asp Gly Gly Arg Arg Lys Lys Gly
210 215 220

Gly Trp Phe Gly Lys His Arg Gly Gln Gly Gly Ser Asn Pro Lys Phe
225 230 235 240

Glu Asn Ile Ala Glu Gly Leu Arg Ala Leu Leu Ala Arg Ser His Val
245 250 255

Glu Arg Thr Thr Asp Glu Gly Thr Trp Val Ala Gly Val Phe Val Tyr
260 265 270

Gly Gly Ser Lys Thr Ser Leu Tyr Asn Leu Arg Arg Gly Thr Ala Leu
275 280 285

Ala Ile Pro Gln Cys Arg Leu Thr Pro Leu Ser Arg Leu Pro Phe Gly
290 295 300

Met Ala Pro Gly Pro Gly Pro Gln Pro Gly Pro Leu Arg Glu Ser Ile
305 310 315 320

Val Cys Tyr Phe Met Val Phe Leu Gln Thr His Ile Phe Ala Glu Val
325 330 335

```
Leu Lys Asp Ala Ile Lys Asp Leu Val Met Thr Lys Pro Ala Pro Thr
            340                 345                 350

Cys Asn Ile Arg Val Thr Val Cys Ser Phe Asp Asp Gly Val Asp Leu
        355                 360                 365

Pro Pro Trp Phe Pro Pro Met Val Glu Gly Ala Ala Ala Glu Gly Asp
    370                 375                 380

Asp Gly Asp Asp Gly Asp Glu Gly Gly Asp Gly Asp Glu Gly Glu Glu
385                 390                 395                     400

Gly Gln Glu Xaa
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 624 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
ATGGGAGAAG GCCCAAGCAC TGGACCCCGG GGTCAGGGTG ATGGAGGCAG GCGCAAAAAA    60
GGAGGGTGGT TTGGAAAGCA TCGTGGTCAA GGAGGTTCCA ACCCGAAATT TGAGAACATT   120
GCAGAAGGTT TAAGAGCTCT CCTGGCTAGG AGTCACGTAG AAAGGACTAC CGACGAAGGA   180
ACTTGGGTCG CCGGTGTGTT CGTATATGGA GGTAGTAAGA CCTCCCTTTA CAACCTAAGG   240
CGAGGAACTG CCCTTGCTAT TCCACAATGT CGTCTTACAC CATTGAGTCG TCTCCCCTTT   300
GGAATGGCCC CTGGACCCGG CCCACAACCT GGCCCGCTAA GGGAGTCCAT TGTCTGTTAT   360
TTCATGGTCT TTTTACAAAC TCATATATTT GCTGAGGTTT TGAAGGATGC GATTAAGGAC   420
CTTGTTATGA CAAAGCCCGC TCCTACCTGC AATATCAGGG TGACTGTGTG CAGCTTTGAC   480
GATGGAGTAG ATTTGCCTCC CTGGTTTCCA CCTATGGTGG AAGGGGCTGC CGCGGAGGGT   540
GATGACGGAG ATGACGGAGA TGAAGGAGGT GATGGAGATG AGGGTGAGGA AGGGCAGGAG   600
CTGCGTCGTG CTTCTGTTGG TTAA                                          624
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 208 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: unknown

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met Gly Glu Gly Pro Ser Thr Gly Pro Arg Gly Gln Gly Asp Gly Gly
1               5                   10                  15

Arg Arg Lys Lys Gly Gly Trp Phe Gly Lys His Arg Gly Gln Gly Gly
            20                  25                  30

Ser Asn Pro Lys Phe Glu Asn Ile Ala Glu Gly Leu Arg Ala Leu Leu
        35                  40                  45

Ala Arg Ser His Val Glu Arg Thr Thr Asp Glu Gly Thr Trp Val Ala
    50                  55                  60

Gly Val Phe Val Tyr Gly Gly Ser Lys Thr Ser Leu Tyr Asn Leu Arg
65                  70                  75                  80

Arg Gly Thr Ala Leu Ala Ile Pro Gln Cys Arg Leu Thr Pro Leu Ser
                85                  90                  95

Arg Leu Pro Phe Gly Met Ala Pro Gly Pro Gly Pro Gln Pro Gly Pro
            100                 105                 110

Leu Arg Glu Ser Ile Val Cys Tyr Phe Met Val Phe Leu Gln Thr His
        115                 120                 125

Ile Phe Ala Glu Val Leu Lys Asp Ala Ile Lys Asp Leu Val Met Thr
    130                 135                 140

Lys Pro Ala Pro Thr Cys Asn Ile Arg Val Thr Val Cys Ser Phe Asp
145                 150                 155                 160

Asp Gly Val Asp Leu Pro Pro Trp Phe Pro Pro Met Val Glu Gly Ala
                165                 170                 175

Ala Ala Glu Gly Asp Asp Gly Asp Asp Gly Asp Glu Gly Gly Asp Gly
            180                 185                 190

Asp Glu Gly Glu Glu Gly Gln Glu Leu Arg Arg Ala Ser Val Gly Xaa
        195                 200                 205
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGGAAGCATA TGCTACCC 18

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGGTAGCATA TGCATATGCT TCCC 24

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

AAAAACGTCC AAGTTGCATT G 21

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GATCGGCATA TGGGAGAAGG CCCAAGCACT GGA 33

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CTGGTGGATC CTTAACCAAC AGAAGCACGA CGCAGCTCCT GCCCTTCCTC AC 52

WHAT IS CLAIMED IS:

1. A process for recovering isolated EBNA1 protein or polypeptide comprising:

providing cells having a nucleus containing expressed EBNA1 protein or polypeptide;

recovering the nucleus containing expressed EBNA1 protein or polypeptide from the cells;

separating the nucleus containing expressed EBNA1 protein or polypeptide into a liquid fraction containing the expressed EBNA1 protein or polypeptide and a solid fraction containing substantially all DNA from the nucleus;

separating the liquid fraction from the solid fraction; and

recovering EBNA1 protein or polypeptide from the liquid fraction.

2. A process according to Claim 1, wherein said separating the nucleus is by centrifugation where the liquid fraction is a supernatant and the solid fraction is a pellet.

3. A process according to Claim 1, wherein said recovering EBNA1 protein or polypeptide from the liquid fraction comprises:

subjecting the liquid fraction to a first ammonium sulfate treatment at an ammonium sulfate concentration which forms a solid phase containing contaminant proteins and a liquid phase containing EBNA1 protein or polypeptide;

subjecting the liquid phase containing EBNA1 protein or polypeptide to a second ammonium sulfate treatment at an ammonium sulfate concentration which forms a solid phase containing EBNA1 protein or polypeptide and a liquid phase containing contaminant proteins; and

separating the solid phase containing EBNA1 protein or polypeptide and the liquid phase containing contaminant proteins.

4. A process according to Claim 3, wherein the first ammonium sulfate treatment is at a >0 to 30% ammonium sulfate concentration.

5. A process according to Claim 3, wherein the second ammonium sulfate treatment is at a 30 to 45% ammonium sulfate concentration.

6. A process according to Claim 3, wherein said recovering EBNA1 protein or polypeptide further comprises:

purifying the solid phase containing EBNA1 protein or polypeptide, after said separating, by affinity column chromatography.

7. A process according to Claim 6, wherein the affinity column chromatography is agarose-heparin affinity column chromatography.

8. A process according to Claim 6, wherein the affinity column chromatography is oligonucleotide affinity column chromatography.

9. A process according to Claim 1, wherein said cells are insect cells.

10. A process according to Claim 9, wherein said insect cells are Sf-9 insect cells.

11. A process according to Claim 10, wherein said EBNA1 protein or polypeptide has an amino acid sequence corresponding to SEQ. ID. No. 2.

12. A process according to Claim 10, wherein said EBNA1 protein or polypeptide is encoded by a nucleotide sequence corresponding to SEQ. ID. No. 1.

13. A process according to Claim 2, wherein the supernatant contains less than 5% of DNA.

14. A process according to Claim 1, wherein said EBNA1 protein or polypeptide has an amino acid sequence corresponding to SEQ. ID. No 2.

15. A process according to Claim 1, wherein said EBNA1 protein or polypeptide is encoded by a nucleotide sequence corresponding to SEQ. ID. No. 1.

16. An isolated EBNA1 protein or polypeptide produced by the process of Claim 1.

17. An isolated EBNA1 protein or polypeptide produced by the process of Claim 2.

18. An isolated EBNA1 protein or polypeptide produced by the process of Claim 3.

19. An isolated EBNA1 protein or polypeptide formulation having substantially no components which generate false positive readings when used to detect Epstein-Barr virus in human serum.

20. An isolated EBNA1 protein or polypeptide according to Claim 19, wherein naturally-occurring EBNA1 protein or polypeptide spans a Gly-Ala repeat amino acid sequence and said isolated EBNA1 protein or polypeptide includes no more than 90% of the Gly-Ala repeat amino acid sequence.

21. An isolated EBNA1 protein or polypeptide according to Claim 19, wherein the EBNA1 protein or polypeptide has an amino acid sequence corresponding to SEQ. ID. No. 2.

22. An isolated EBNA1 protein or polypeptide according to Claim 19, wherein the EBNA1 protein or polypeptide is encoded by a nucleotide sequence corresponding to SEQ. ID. No. 1.

23. An isolated EBNA1 protein or polypeptide according to Claim 19, wherein the EBNA1 protein or polypeptide is recombinant.

24. An isolated EBNA1 protein or polypeptide according to Claim 19, wherein the EBNA1 protein or polypeptide is purified.

25. An isolated DNA molecule encoding EBNA1 protein or polypeptide containing substantially no components which generate false positive readings when used to detect Epstein-Barr virus in human serum.

26. An isolated DNA molecule according to Claim 25, wherein said DNA molecule is isolated from any other DNA molecule which expresses protein that generates false positive readings when the EBNA1 protein or polypeptide is used to detect Epstein-Barr virus in human serum.

27. An isolated DNA molecule according to Claim 25, wherein said DNA molecule encodes a protein having an amino acid sequence corresponding to SEQ. ID. No.2.

28. An isolated DNA molecule according to Claim 25, wherein said DNA molecule contains a nucleotide sequence corresponding to SEQ. ID. No. 1.

29. A recombinant DNA expression system comprising an expression vector into which is inserted a heterologous DNA molecule encoding EBNA1 protein or polypeptide containing substantially no components which generate false positive readings when used to detect Epstein-Barr virus in human serum.

30. A recombinant DNA expression system according to Claim 29, wherein said DNA molecule encodes a protein having an amino acid sequence corresponding to SEQ. ID. No. 2.

31. A host cell incorporating a heterologous DNA molecule encoding EBNA1 protein or polypeptide containing substantially no components which generate false positive readings when used to detect Epstein-Barr virus in human serum.

32. A host cell according to Claim 31, wherein said DNA molecule encodes a protein having an amino acid sequence corresponding to SEQ. ID. No. 2.

33. A host cell according to Claim 31, wherein said host cell is an insect cell.

34. A host cell according to Claim 33, wherein said insect cell is an Sf-9 insect cell.

35. A method for detection of Epstein-Barr virus in a sample of human tissue or body fluids comprising:

providing an isolated EBNA1 protein or polypeptide formulation according to Claim 20 as an antigen;

contacting the sample with the antigen; and

detecting any reaction which indicates that Epstein-Barr virus is present in the sample using an assay system.

36. A method according to Claim 35, wherein said assay system is selected from the group consisting of an enzyme-linked immunosorbant assay, a radioimmunoassay, a gel diffusion precipitin reaction assay, an immunodiffusion assay, an agglutination assay, a fluorescent immunoassay, a protein A immunoassay, and an immunoelectrophoresis assay.

37. A method according to Claim 36, wherein said isolated EBNA1 protein or polypeptide is produced according to Claim 1.

38. A method of expressing an EBNA1 protein coding sequence in a cell, wherein said method comprises the steps of:

cloning an EBNA1 protein coding sequence into a baculovirus transfer vector;

co-transfecting insect cells with said baculovirus transfer vector and *Autographica californica* nuclear polyhedrosis genomic DNA;

recovering recombinant baculoviruses; and

infecting cells with said recombinant baculovirus under conditions facilitating expression of isolated EBNA1 protein or polypeptide in the cell,

wherein naturally-occurring EBNA1 protein coding sequence spans a Gly-Ala repeat amino acid sequence and said EBNA1 protein coding sequence includes no less than 90% of the Gly-Ala repeat amino acid sequence.

39. A method according to Claim 38, wherein said insect cells are Sf-9 insect cells.

40. A method according to Claim 38, wherein said EBNA1 protein coding sequence has an amino acid sequence corresponding to SEQ. ID. No. 2.

Figure 1

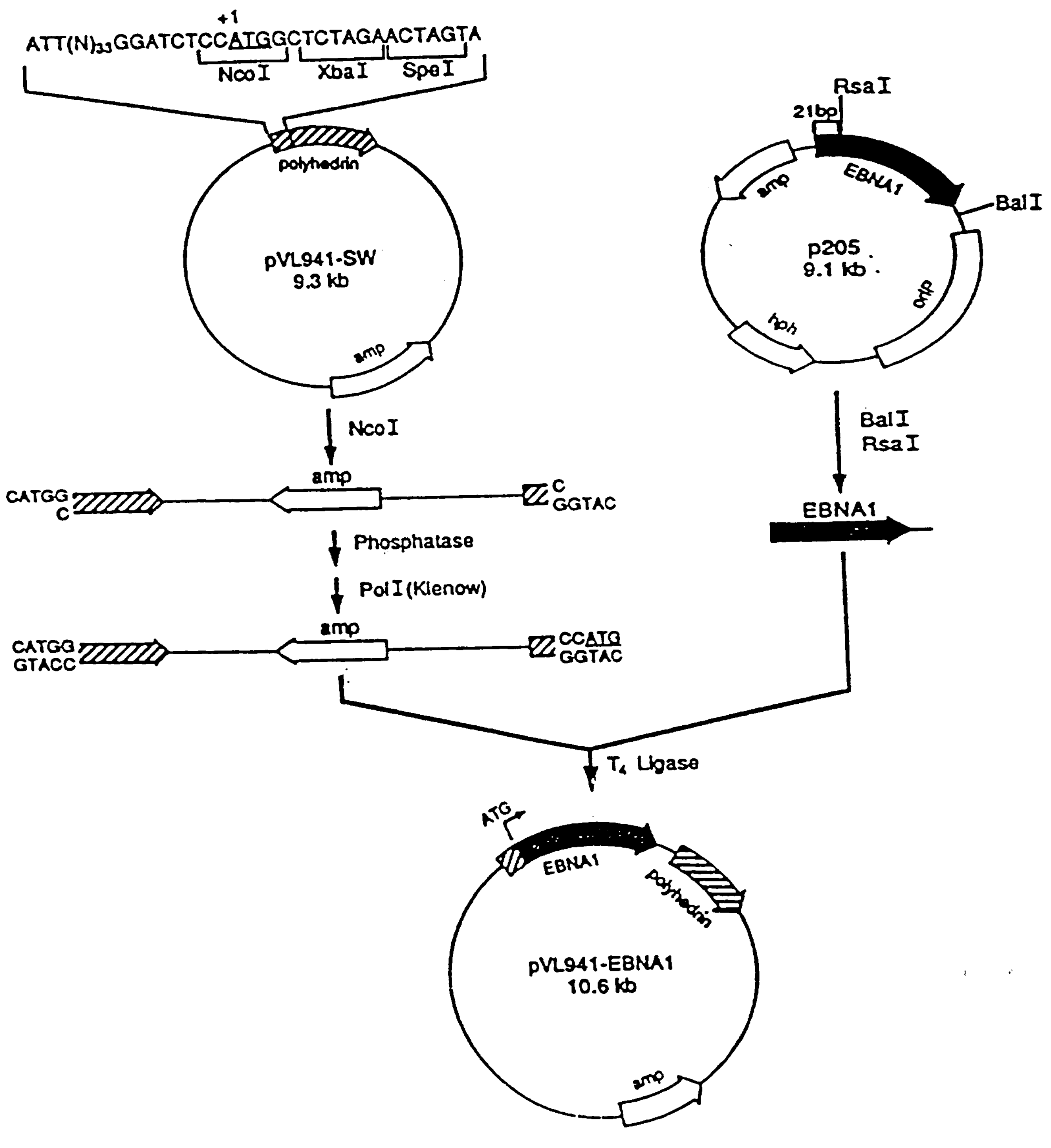

+1
ATT(N)33GGATCTCCATGGCTCTAGAACTAGTA
NcoI
XbaI
SpeI
polyhedrin
pVL941-SW
9.3 kb
amp
RsaI
21bp
EBNA1
amp
BalI
p205
9.1 kb
oriP
hph
NcoI
amp
CATGG
C
C
GGTAC
BalI
RsaI
EBNA1
Phosphatase
Pol I (Klenow)
amp
CATGG
GTACC
CCATG
GGTAC
T4 Ligase
ATG
EBNA1
polyhedrin
pVL941-EBNA1
10.6 kb
amp

Figure 2

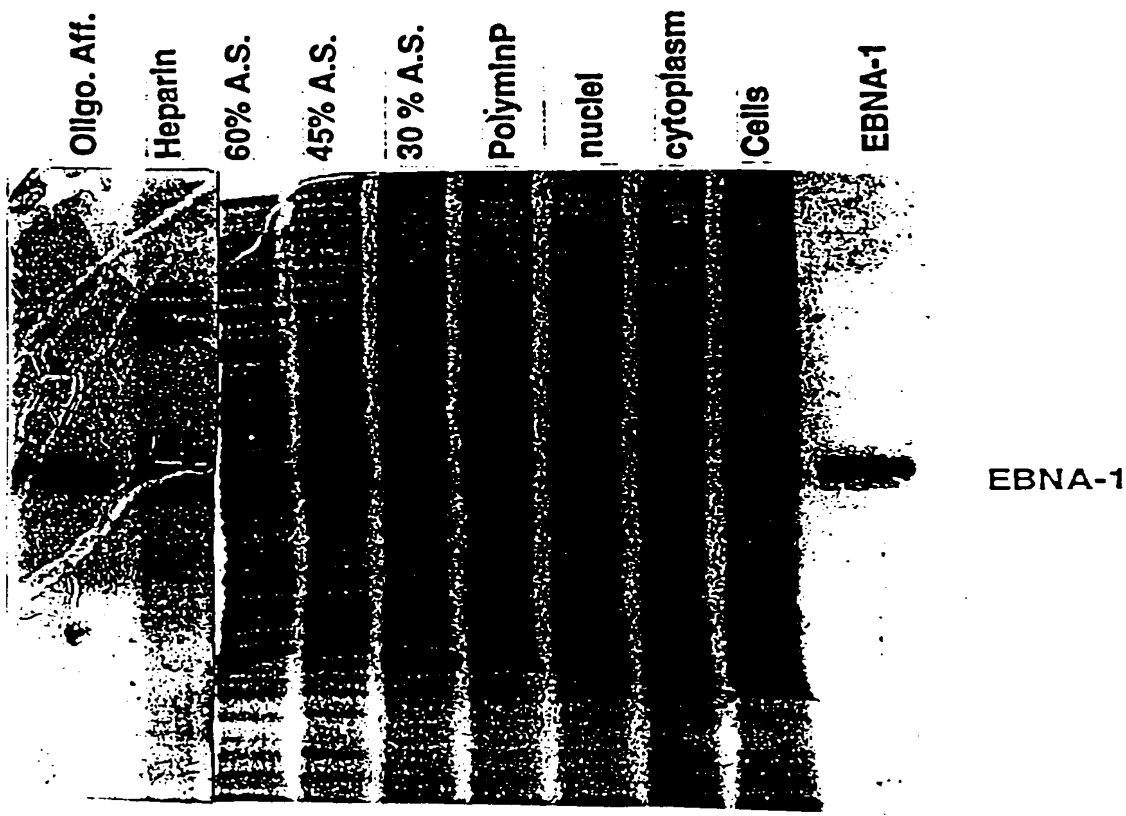

Oligo. Aff.
Heparin
60% A.S.
45% A.S.
30 % A.S.
PolyminP
nuclei
cytoplasm
Cells
EBNA-1
EBNA-1

Figure 3

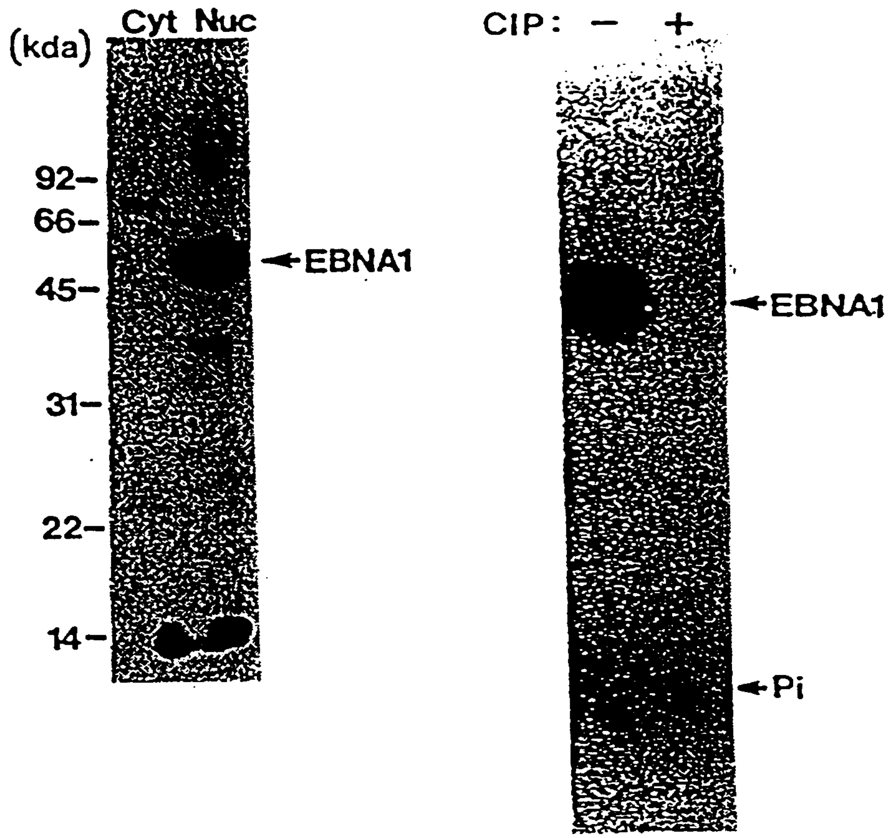


Figure 4

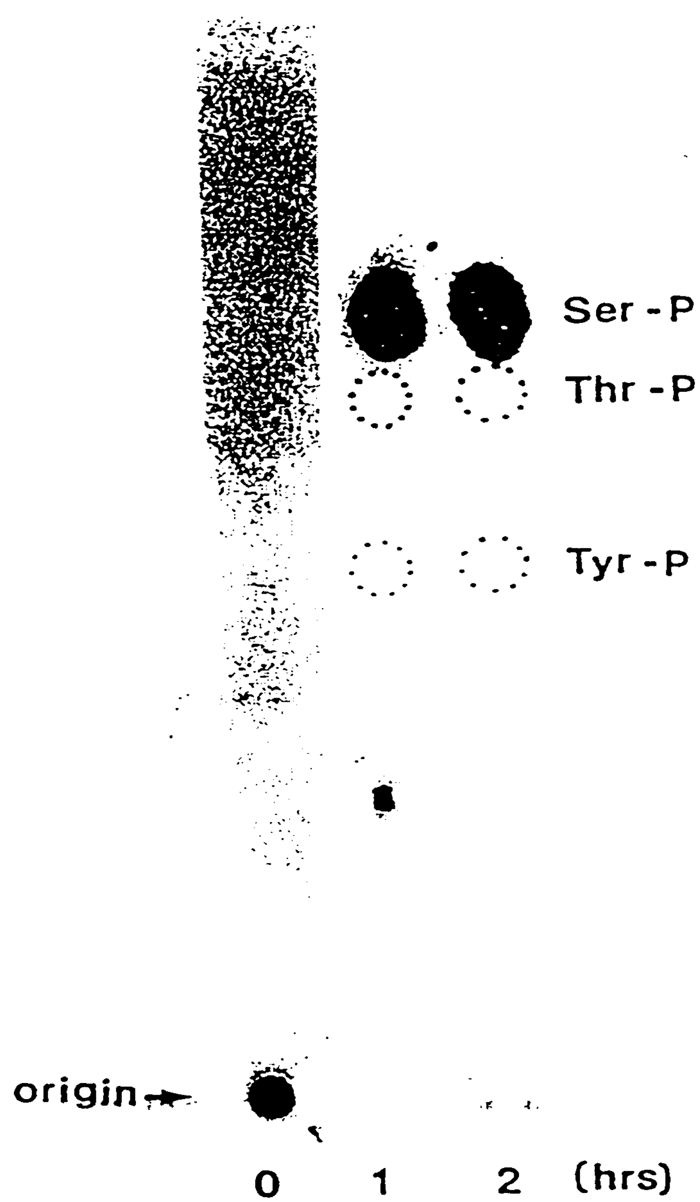


Figure 5A

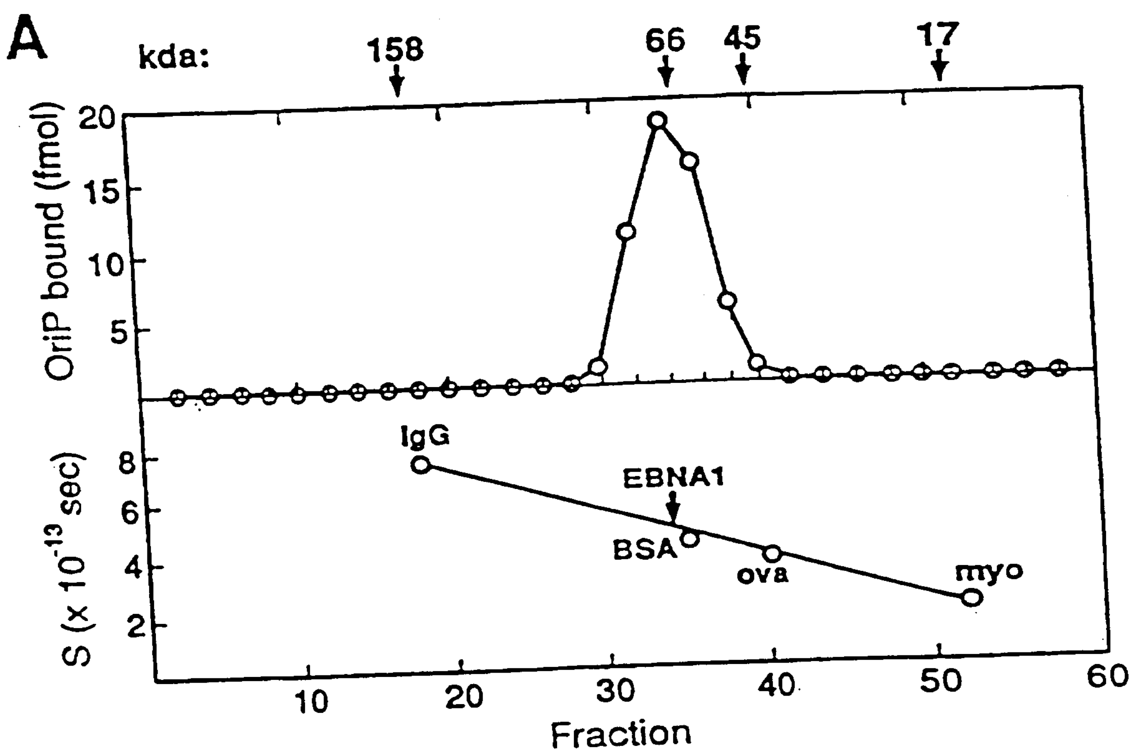

A
kda:
158
66
45
17
OriP bound (fmol)
20
15
10
5
S (x $10^{-13}$ sec)
8
6
4
2
IgG
EBNA1
BSA
ova
myo
10
20
30
40
50
60
Fraction

Figure 5B

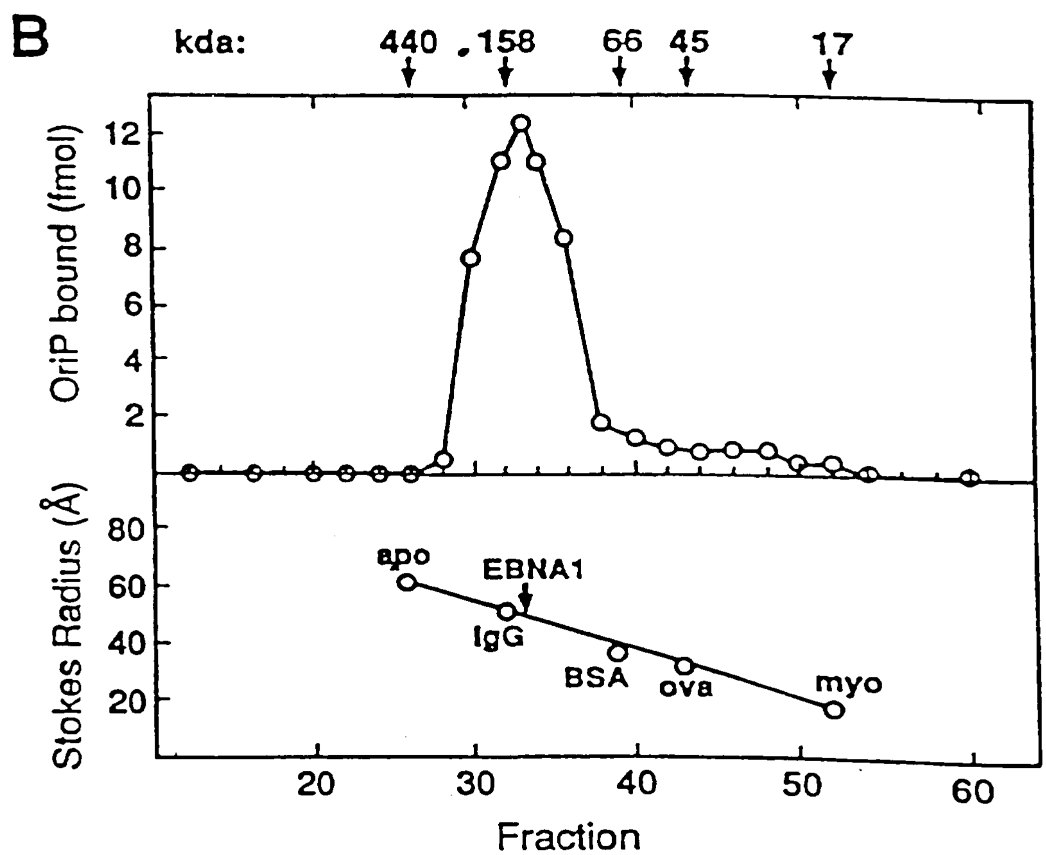

B
kda:
440
158
66
45
17
OriP bound (fmol)
12
10
8
6
4
2
Stokes Radius (Å)
80
60
40
20
apo
EBNA1
IgG
BSA
ova
myo
20
30
40
50
60
Fraction

Figure 6

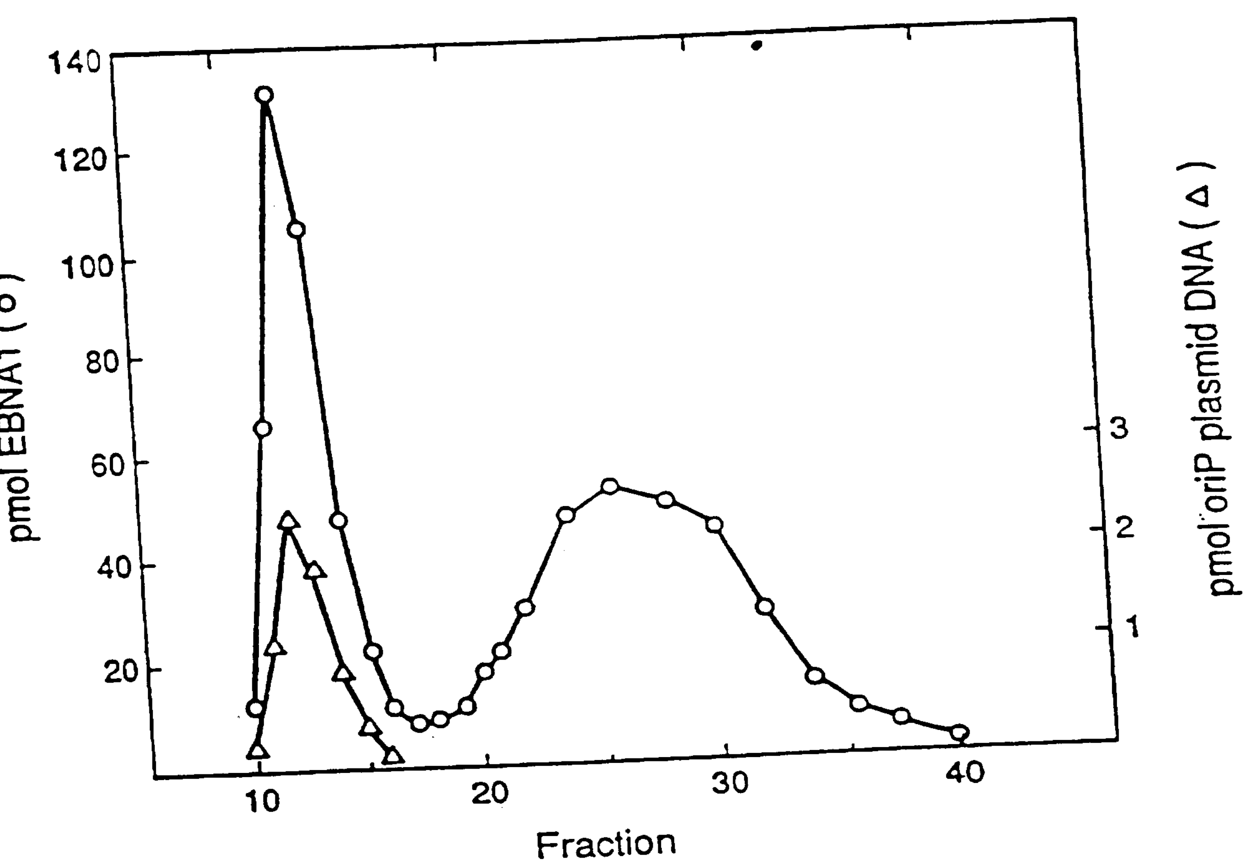

140
120
100
80
60
40
20
pmol EBNA1 ( o )
3
2
1
pmol oriP plasmid DNA ( Δ )
10
20
30
40
Fraction

Figure 7

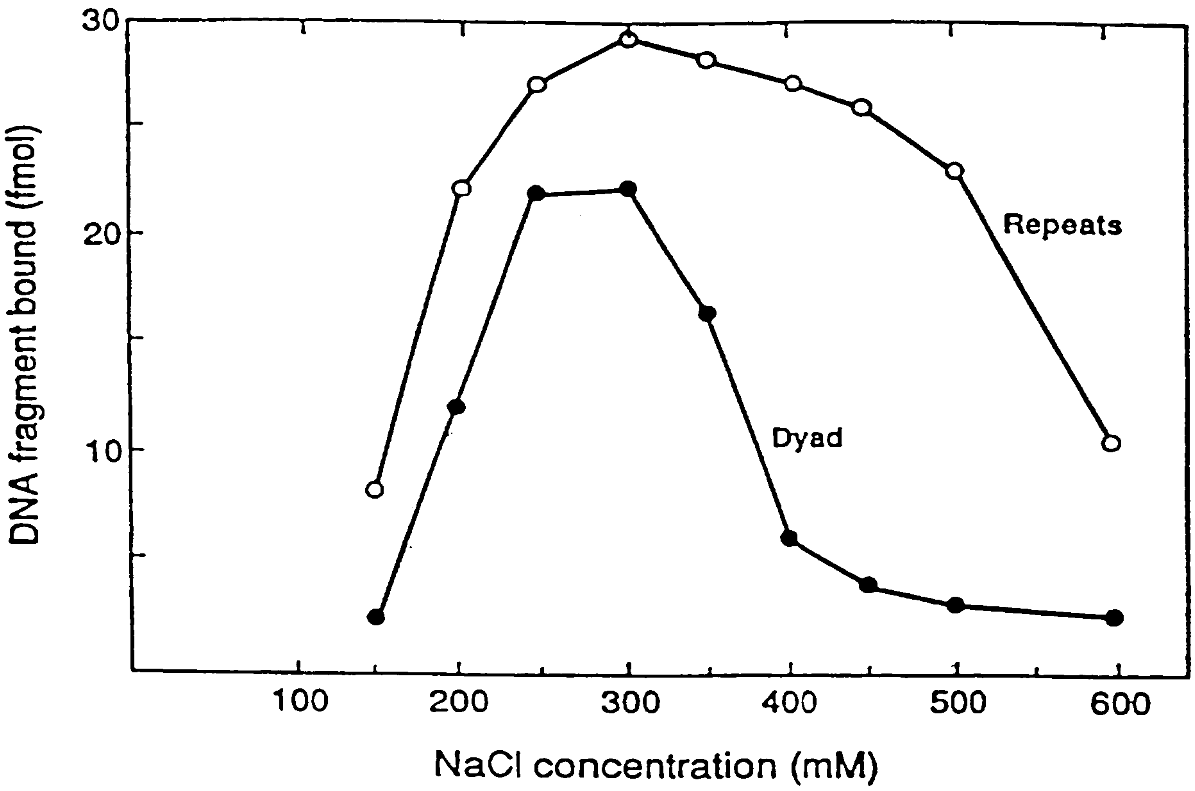

30
20
10
DNA fragment bound (fmol)
100
200
300
400
500
600
NaCl concentration (mM)
Repeats
Dyad

Figure 8

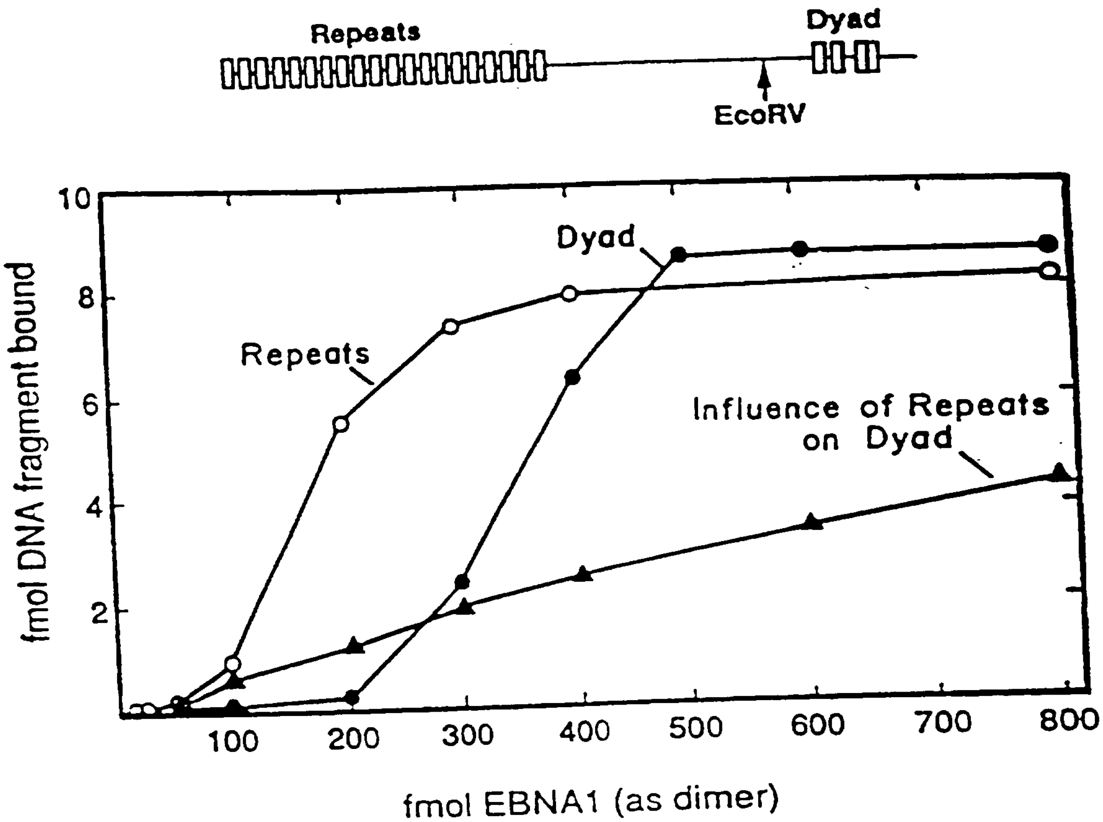

Repeats
Dyad
EcoRV
10
8
6
4
2
100
200
300
400
500
600
700
800
fmol DNA fragment bound
fmol EBNA1 (as dimer)
Dyad
Repeats
Influence of Repeats on Dyad

Figure 9A

A

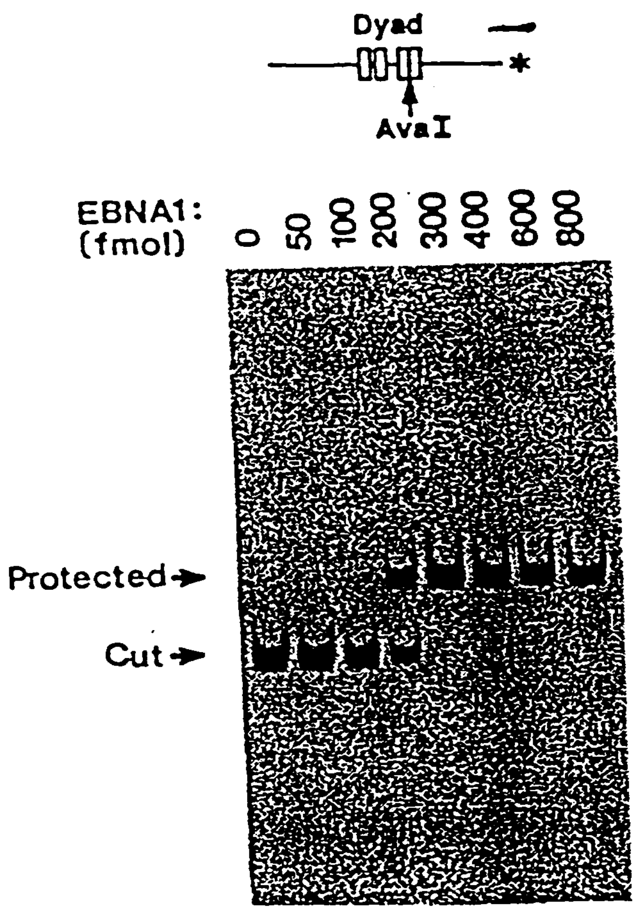

Dyad
AvaI
*
EBNA1:
(fmol)
0
50
100
200
300
400
600
800
Protected
Cut

Figure 9B

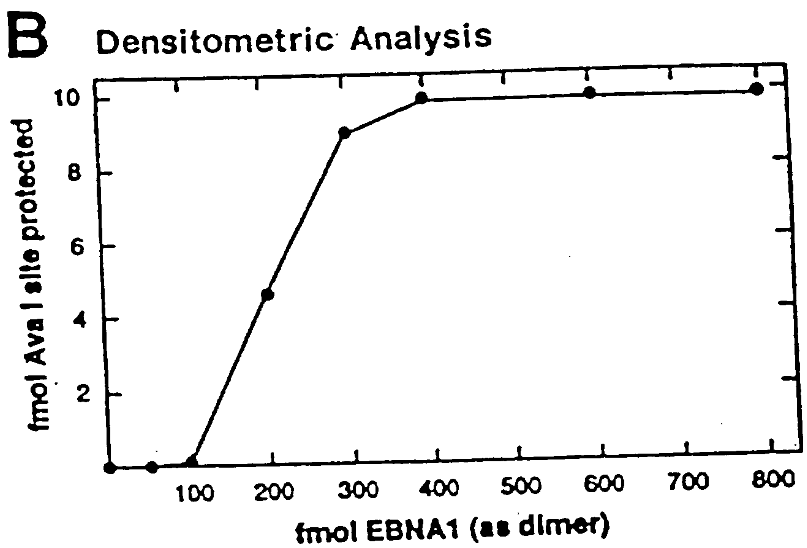

B
Densitometric Analysis
10
8
6
4
2
fmol Ava I site protected
100
200
300
400
500
600
700
800
fmol EBNA1 (as dimer)

Figure 10

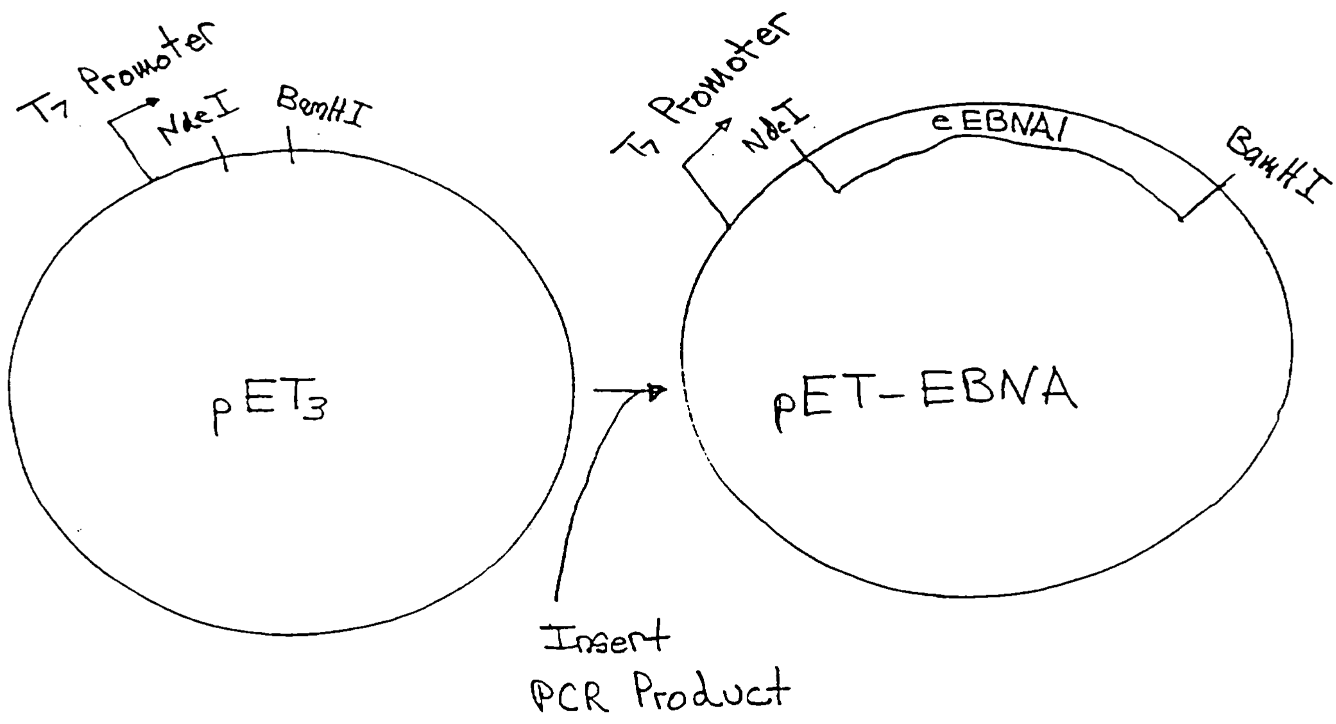
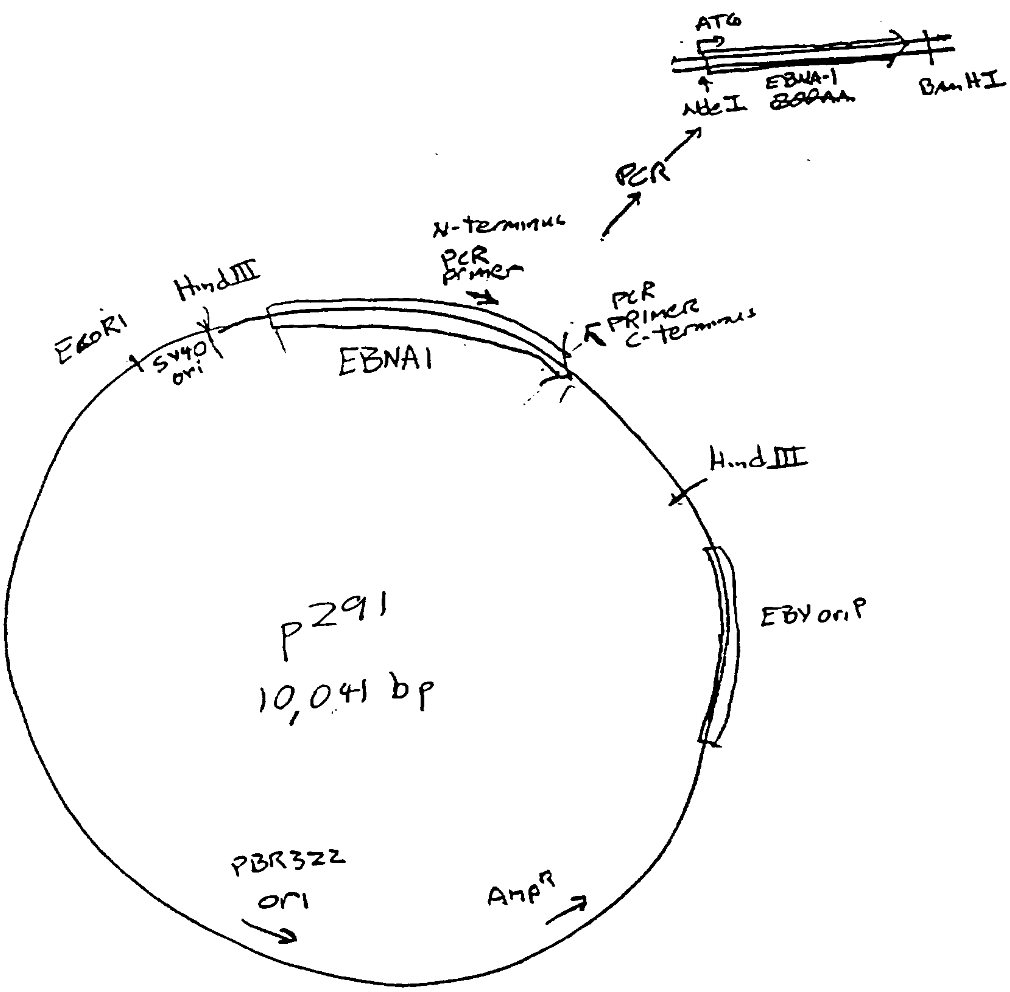

T7 Promoter
NdeI
BamHI
pET3
T7 Promoter
NdeI
eEBNA1
BamHI
pET-EBNA
Insert
PCR Product

Figure 11


ATG
EBNA-1
800AA
NheI
BamHI
PCR
N-terminus
PCR
primer
PCR
PRIMER
C-terminus
HindIII
EcoRI
SV40
ori
EBNA1
HindIII
EBV oriP
p291
10,041 bp
PBR322
ori
AmpR

```
Sau3A I 1; Mbo I 1; Dpn I 1; Nde I 7; Sau96 I 19; Hae III 19; PflM I 22; Tth111 II 23;
Sau96 I 31; Nla IV 31; Ava II 31; Sec I 34; Ava I 35; Bcn I 35; Nci I 35; ScrF I 35; Sec I 35;
Sma I 35; Xma I 35; Bcn I 36; Hpa II 36; Msp I 36; Nci I 36; ScrF I 36; Sec I 36;
Hph I 46; Mnl I 53; Hha I 60; HinP I 60; Mnl I 71

GATCGGCATATGGGAGAAGGCCCAAGCACTGGACCCCGGGGTCAGGGTGATGGAGGCAGGCGCAAAAAAGGAGGGTGGTT   80
CTAGCCGTATACCCTCTTCCGGGTTCGTGACCTGGGGCCCCAGTCCCACTACCTCCGTCCGCGTTTTTTCCTCCCACCAA

SfaN I 87; Mnl I 101; Nla IV 103; Mse I 139; Ban II 143; Bsp1286 I 143; HgiA I 143;
Sac I 143; Alu I 144; BstN I 150; EcoR II 150; ScrF I 150; Mae I 155; Hinf I 159; Ple I 159

TGGAAAGCATCGTGGTCAAGGAGGTTCCAACCCGAAATTTGAGAACATTGCAGAAGGTTTAAGAGCTCTCCTGGCTAGGA   160
ACCTTTCGTAGCACCAGTTCCTCCAAGGTTGGGCTTTAAACTCTTGTAACGTCTTCCAAATTCTCGAGAGGACCGATCCT

Mae III 161; Mae II 164; Cfr10 I 199; Hpa II 200; Msp I 200; Mnl I 218; Mnl I 230

GTCACGTAGAAAGGACTACCGACGAAGGAACTTGGGTCGCCGGTGTGTTCGTATATGGAGGTAGTAAGACCTCCCTTTAC   240
CAGTGCATCTTTCCTGATGGCTGCTTCCTTGAACCCAGCGGCCACACAAGCATATACCTCCATCATTCTGGAGGGAAATG
```

Figure 12A

15/18

ScrF I
EcoR II
BstN I
Sec I
Sau96 I
Nla IV
Hae III
Ple I
Hinf I
Dde I Mnl I
Bsu36 I

AACCTAAGGCGAGGAACTGCCCTTGCTATTCCACAATGTCGTCTTACACCATTGAGTCGTCTCCCCTTTGGAATGGCCCC 320
TTGGATTCCGCTCCTTGACGGGAACGATAAGGTGTTACAGCAGAATGTGGTAACTCAGCAGAGGGGAAACCTTACCGGGG

243
244
251
294
294
315
315
315
318
319
319
319

Sau96 I
Hae III
Msp I
Hpa II
ScrF I
Nci I
Bcn I
Sau96 I
Nla IV
Ava II
PflM I
Sau96 I
ScrF I
EcoR II
BstN I
Hae III
Dde I
Ple I
Hinf I
Nla III

TGGACCCGGCCCACAACCTGGCCCGCTAAGGGAGTCCATTGTCTGTTATTTCATGGTCTTTTTACAAACTCATATATTTG 400
ACCTGGGCCGGGTGTTGGACCGGGCGATTCCCTCAGGTAACAGACAATAAAGTACCAGAAAAATGTTTGAGTATATAAAC

322
322
322
325
325
325
326
326
328
328
331
337
337
337
340
340
346
352
352
372

Sau96 I
Ava II
PpuM I
Eco0109 I
Mnl I
Dde I
SfaN I
Fok I
Mse I
BspM I
Mae III
Hph I
Fnu4H I
Bbv I

CTGAGGTTTTGAAGGATGCGATTAAGGACCTTGTTATGACAAAGCCCGCTCCTACCTGCAATATCAGGGTGACTGTGTGC 480
GACTCCAAAACTTCCTACGCTAATTCCTGGAACAATACTGTTTCGGGCGAGGATGGACGTTATAGTCCCACTGACACACG

401
403
414
415
422
425
425
426
426
454
468
469
479
479

BstU I
Sec I
Sac II
NspB II
Fnu4H I
ScrF I
EcoR II
BstN I
Sec I
Mnl I
Alu I
Fnu4H I
Bbv I
Mnl I
Hph I

AGCTTTGACGATGGAGTAGATTTGCCTCCCTGGTTTCCACCTATGGTGGAAGGGGCTGCCGCGGAGGGTGATGACGGAGA 560
TCGAAACTGCTACCTCATCTAAACGGAGGGACCAAAGGTGGATACCACCTTCCCCGACGGCGCCTCCCACTACTGCCTCT

481
505
508
509
509
509
535
535
538
539
539
539
540
544
547

```
                                                                                  Sau3A I
                                                                                  Mbo I
                                                                                  Dpn I
                                                                                  Alw I
                                                                                 Nla IV
                                                  Fnu4H I                        BstY I
                                            Mnl I Bbv I                          BamH I
              Hph I           Hph I              Alu I                           Alw I
            Mnl I          Mnl I            AlwN I  Hga I                    Mse I
            | |            |  |  |          |   ||  |                        |   ||
TGACGGAGATGAAGGAGGTGATGGAGATGAGGGTGAGGAAGGGCAGGAGCTGCGTCGTGCTTCTGTTGGTTAAGGATCCA   640
ACTGCCTCTACTTCCTCCACTACCTCTACTCCCACTCCTTCCCGTCCTCGACGCAGCACGAAGACAACCAATTCCTAGGT
            | |            |  |  |          |   ||  |                        |   ||

            575            589              604     612                      630
              577             592               608                              634
                                 595             609                             634
                                                 609                             634
                                                                                 634
                                                                                  635
                                                                                  635
                                                                                  635
                                                                                  635

CCAG   644
GGTC
```

Figure 12C

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/08700

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) :C07K 1/00; C12Q 1/70; C12P 21/06; C12N 5/00; C07H 21/04
US CL :530/350, 413, 418; 435/5, 69.3, 240.2, 320.1; 536/23.72
According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 530/350, 413, 418; 435/5, 69.3, 240.2, 320.1; 536/23.72

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Dialog

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | The Journal of Biological Chemistry, Vol. 253, No. 8, issued 25 April 1978, D. Baron et al., "Partial Purification and Properties of the Epstein-Barr Virus-associated Nuclear Antigen," pages 2875-2881, see entire document. | 1-8, 11, 13-28, 35, 36 |
| Y | Journal of Medical Virology, Vol. 39, issued 1993, A. Hille et al., "Expression of Epstein-Barr Virus Nuclear Antigen 1, 2A, 2B in the Baculovirus Expression System: Serological Evaluation of Human Antibodies to These Proteins," pages 233-241, see entire document. | 1-36, 38-40 |

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be part of particular relevance
- "E" earlier document published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed
- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 17 OCTOBER 1995 | 01 NOV 1995 |

| Name and mailing address of the ISA/US | Authorized officer |
|---|---|
| Commissioner of Patents and Trademarks<br>Box PCT<br>Washington, D.C. 20231<br>Facsimile No. (703) 305-3230 | DONNA C. WORTMAN, PH.D.<br>Telephone No. (703) 308-0196 |

Form PCT/ISA/210 (second sheet)(July 1992)*

| International application No. |
|---|
| PCT/US95/08700 |

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | Journal of Virological Methods, Vol. 42, issued 1993, I. Farber et al., "Serological diagnosis of infectious mononucleosis using three anti-Epstein-Barr virus recombinant ELISAs," pages 301-308, see entire document. | 1-36, 38-40 |
| Y | International Journal of Cancer, Volume 22, issued 1978, T. Matsuo et al., "Studies on Epstein-Barr Virus-Related Antigens. III. Purification of the Virus-Determined Nuclear Antigen (EBNA) from Non-Producer Raji Cells," pages 747-752, see entire document. | 1-36, 38-40 |
| Y | Journal of Cellular Biochemistry, Supplement 13D, issued 1989, L. Frappier et al., "Purification of the Epstein-Barr Virus Origin Binding Protein (EBNA 1) from an E. coli Overproducing Strain," page 154, see Abstract L419. | 1-36, 38-40 |
| Y | Methods in Enzymology, Volume 182, published 1990, Academic Press, San Diego, M.P. Deutscher, Ed., pages 285-300, especially pages 293-296. | 1-25, 35, 36 |
| Y | WO, A, 94/06912 (AKZO N.V.) 31 March 1994, see entire document. | 1-36, 38-40 |
| Y | US, A, 5,256,768 (MILMAN) 26 October 1993, see entire document. | 1-36, 38-40 |

**INTERNATIONAL SEARCH REPORT**

International application No.
PCT/US95/08700

**Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)**

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 37
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)**

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.